```
TITLE      National Institutes of Health, [illegible]
JOURNAL    Unpublished (1999)
COMMENT    [illegible]
           Email: [illegible]
           [illegible]
           cDNA Library Preparation: [illegible] Laboratory
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium [illegible]
           DNA Sequencing by: [illegible]
           Clone distribution: [illegible]
           [illegible] through the I.M.A.G.E. Consortium/[illegible]
           http://image.llnl.gov
           [illegible]
           High quality sequence stops [illegible]
FEATURES            Location/Qualifiers
     source         [illegible]
                    /organism="Homo sapiens"
                    /db_xref="[illegible]"
                    /clone="IMAGE:[illegible]"
                    /clone_lib="NIH_MGC_[illegible]"
                    /tissue_type="epidermoid carcinoma, cell line"
                    /lab_host="[illegible] (phage resistant)"
                    /note="Organ: [illegible]; Vector: [illegible]; Site_1: [illegible]
                    Site_2: [illegible] cDNA made [illegible]
                    directionally cloned into [illegible] sites [illegible]
                    following [illegible] adapter [illegible]. Library [illegible]
                    by [illegible] in the laboratory of Gerald M. [illegible]
                    (University of California, Berkeley) using [illegible] cDNA
                    synthesis kit (Stratagene) and [illegible]
                    Technologies). Note: this is a NIH_MGC Library."
BASE COUNT    [illegible]
ORIGIN
```

[illegible]

[illegible]

RESULT 3
[illegible]
LOCUS [illegible] mRNA linear EST [illegible] 2002
DEFINITION AGENCOURT [illegible] NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6274622 5', mRNA sequence.
ACCESSION [illegible]
VERSION [illegible] GI:[illegible]
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to [illegible])
AUTHORS NIH MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLCM2456 row: [illegible]
High quality sequence start: 1
High quality sequence stop: 624.
FEATURES Location/Qualifiers
source 1..[illegible]
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6274622"
/clone_lib="NIH_MGC_102"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT primer. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: [illegible]. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP cDNA Synthesis Kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
BASE COUNT [illegible]
ORIGIN

Alignment Scores:
Pred. No.: [illegible] Length: [illegible]
Score: [illegible] Matches: [illegible]
Percent Similarity: [illegible] Conservative: [illegible]
Best Local Similarity: [illegible] Mismatches: [illegible]
Query Match: [illegible] Indels: [illegible]
DB: [illegible] Gaps: [illegible]

[illegible]

RESULT 4
[illegible]
LOCUS [illegible] mRNA linear EST [illegible]
DEFINITION AGENCOURT [illegible] NIH_MGC_102 Homo sapiens cDNA clone IMAGE:[illegible] 5', mRNA sequence.
ACCESSION [illegible]
VERSION [illegible] GI:[illegible]
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to [illegible])
AUTHORS NIH MGC http://mgc.nci.nih.gov/.

[illegible]

```
RESULT 2
BM55[illegible]
LOCUS       BM5[illegible]     1042 bp    mRNA    linear   EST [illegible]
DEFINITION  AGENC[illegible] NIH_MGC_41 Homo sapiens cDNA clone IMAGE:[illegible]
            5' mRNA sequence.
ACCESSION   BM55[illegible]
VERSION     BM55[illegible]  GI:[illegible]
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1042)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            [illegible]
            Tissue Procurement: [illegible]
            cDNA Library Preparation: [illegible]
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: [illegible]
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: [illegible]
            High quality sequence stop: [illegible]
FEATURES             Location/Qualifiers
     source          1..1042
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:[illegible]"
                     /clone_lib="NIH_MGC_41"
                     /tissue_type="[illegible]"
                     /lab_host="[illegible] (phage resistant)"
                     /note="[illegible]
                     [illegible]
                     into EcoRI/XhoI sites using the following [illegible]
                     [illegible]
                     [illegible]
                     Berkeley) using ZAP-cDNA [illegible]
                     Superscript II RT [illegible]
                     NIH MGC Library."
BASE COUNT  [illegible]
```

RESULT [illegible]

Alignment Scores:
[illegible]

REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. [illegible], [illegible]-44 (1999)
MEDLINE [illegible]
PUBMED [illegible]
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., [illegible],Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 1[illegible] (1[illegible]), [illegible] (2000)
MEDLINE 20499374
PUBMED [illegible]
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., [illegible], [illegible], Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 1[illegible] (11), [illegible] (2000)
MEDLINE [illegible]
PUBMED [illegible]
REFERENCE 4
AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., [illegible], Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashizaki,Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 4[illegible] (6821), [illegible] (2001)
MEDLINE 2[illegible]
PUBMED [illegible]
REFERENCE 5 (bases 1 to 3[illegible])
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Butler,[illegible], Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Gariboldi,M., Hara,A., Hayatsu,N., Hill,D., Hirozane,K., Hiraoka,T., [illegible], Ikegami,T., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., [illegible], Ozawa,K., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted ([illegible]) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, [illegible] Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa [illegible], Japan (E-mail:[illegible]; URL:http://genome.gsc.riken.go.jp/, [illegible])

COMMENT [illegible]

FEATURES [illegible]

BASE COUNT [illegible]
ORIGIN

SUMMARIES

ALIGNMENTS

LOCUS mRNA linear

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE RIKEN full-length enriched mouse cDNA library

ORGANISM Mus musculus

[illegible]

```
KEYWORDS
ORGANISM
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
            Mammalia; [illegible] Rodentia; [illegible]
REFERENCE   [illegible]
  AUTHORS   [illegible] Woodruff [illegible] Sutherland [illegible]
  TITLE     [illegible]
            dipeptidyl peptidase (DPP) [illegible] DPP8
  JOURNAL   [illegible]
  MEDLINE   [illegible]
REFERENCE   [illegible]
  AUTHORS   [illegible] McCaughan,G.W. and Gorrell,M.D.
  TITLE     [illegible]
  JOURNAL   [illegible] A.W. Morrow [illegible]
            Centenary Institute of Cell Biology and Cancer [illegible]
            [illegible] Newtown, Sydney, NSW 2042, Australia
FEATURES             [illegible]
                     /note="serine protease; member of [illegible]
                     [illegible]
                     [illegible] this mRNA [illegible] GenBank [illegible]
                     Numbers AF221634, AF221635 and AF221[illegible]"
                     /product="dipeptidyl peptidase 8"
                     /protein_id="AA[illegible]"
                     /translation="[illegible]YWW[illegible]
                     [illegible]ANEKVIEKMSLIM[illegible]
                     [illegible]YVVSYVN[illegible]
                     [illegible]YKLSSP[illegible]CKIKEEWA[illegible]
                     [illegible]KKYF[illegible]VNNF[illegible]
                     [illegible]
                     [illegible]VAIAGAPVIWIFY[illegible]YI"
BASE COUNT
ORIGIN
```

[illegible]

```
Qy   [illegible]
Db   [illegible]
Qy   [illegible] AlaLeuLysValIle 982
Db   [illegible] AAAA[illegible]ATA [illegible]

RESULT 14
HSM8[illegible]
LOCUS       HSM8[illegible]     [illegible] bp    mRNA      linear   [illegible]
DEFINITION  Homo sapiens mRNA; cDNA DKFZp[illegible]M[illegible] (from clone DKFZp[illegible]
ACCESSION   AL[illegible]
VERSION     AL[illegible]  GI:[illegible]
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to [illegible])
  AUTHORS   Ansorge,W., Wirkner,[illegible], Mewes,H.W., Weil,B. and Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted ([illegible]) [illegible] Neuenheim, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
            sequenced by EMBL (European Molecular Biology Laboratory),
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            This clone (DKFZp[illegible]M[illegible]) is available at the RZPD in Berlin.
            Please contact the RZPD, Ressourcenzentrum, Heubnerweg [illegible]
            Berlin-Charlottenburg, GERMANY; [illegible]
            Information about the clone and the sequencing project is [illegible]
            at http://mips.gsf.de/[illegible]/cDNA/.
FEATURES             Location/Qualifiers
     source          1..[illegible]
                     /organism="Homo sapiens"
                     /db_xref="taxon:[illegible]"
                     /map="[illegible]"
                     /clone="DKFZp[illegible]M[illegible]"
                     /tissue_type="melanoma (MeWo cell line)"
                     /clone_lib="[illegible] (synonym: [illegible]). Vector [illegible]
                     DH10B; sites NotI + SalI"
                     /dev_stage="[illegible]"
     gene            1..[illegible]
                     /gene="DKFZp[illegible]M[illegible]"
     CDS             <1..[illegible]
                     /gene="DKFZp[illegible]M[illegible]"
                     /note="unknown protein"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="CAB[illegible]"
                     /db_xref="GI:[illegible]"
                     /translation="[illegible]WVANIF[illegible]SNVL[illegible]
                     FI[illegible]YWW[illegible]ASWE[illegible]KI[illegible]V[illegible]
                     KNPKIALKLALP[illegible]KLL[illegible]KV[illegible]
                     [illegible]
                     FYPFPNQRERISSAPSAPMNAR[illegible]ASA[illegible]KS[illegible]WS[illegible]
                     EP[illegible]S[illegible]EWEVLA[illegible]SK[illegible]WVNE[illegible]
                     IPGFSHSCSMSQNPSMPVGHY[illegible]VHVYKLS[illegible]WASMP[illegible]
                     [illegible]YVPE[illegible]RSLVR[illegible]

BASE COUNT   [illegible]
ORIGIN

Alignment Scores:
Pred. No.:           [illegible]
Score:               [illegible]
Percent Similarity:  [illegible]
Best Local Similarity: [illegible]
Query Match:         [illegible]
DB:                  [illegible]

US-[illegible]

Qy   [illegible]
Db   [illegible]
```

Sydney, N.S.W. 2..., Australia

family belonging to ...

splice variants of this mRNA are found in GenBank

/codon_start=1

/protein_id="AA..."

[illegible]

Alignment Scores:

Pred. No.:	[illegible]	Length:	[illegible]
Score:	2870.00	Matches:	[illegible]
Percent Similarity:	[illegible]	Conservative:	[illegible]
Best Local Similarity:	61.55%	Mismatches:	[illegible]
Query Match:	[illegible]	Indels:	[illegible]
DB:	[illegible]	Gaps:	[illegible]

[illegible]

DEFINITION

KEYWORDS

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS

TITLE

JOURNAL

FEATURES

/organism="Homo sapiens"

[illegible]

```
Qy   982 ValIle 982

Db   [illegible] GTAATA [illegible]

RESULT [illegible]
AK[illegible]
LOCUS       AK[illegible]                  [illegible] bp    mRNA    linear   [illegible]
DEFINITION  Homo sapiens cDNA [illegible]
            [illegible] DIPEPTIDYL PEPTIDASE IV ([illegible]).
ACCESSION   AK[illegible]
VERSION     AK[illegible]
KEYWORDS    [illegible] capping; [illegible] (full insert sequence).
SOURCE      Homo sapiens placenta cDNA to mRNA, clone lib:PLACE1
            clone:PLACE1[illegible].
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   [illegible]
```

[illegible]

[illegible]

Alignment Scores:

Pred. No.:	[illegible]	Length:	[illegible]
Score:	[illegible]	Matches:	[illegible]
Percent Similarity:	[illegible]	Conservative:	[illegible]
Best Local Similarity:	[illegible]	Mismatches:	[illegible]
Query Match:	[illegible]	Indels:	[illegible]
DB:	[illegible]	Gaps:	[illegible]

US-10-070-464-1 (1-882) x AK000200 (1-2774)

[illegible]

[illegible]

REFERENCE [illegible]
AUTHORS [illegible]
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE [illegible]
AUTHORS [illegible] and Nakamura,Y.
TITLE Direct Submission
JOURNAL [illegible] Institute of Medical Science, University of Tokyo, Department of Virology; Shirokanedai, [illegible], Minato-ku, Tokyo 108-8639, Japan (E-mail:[illegible]; Tel:81-3-[illegible], Fax:81-3-5449-5410)
COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing: Department of Virology, Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).
FEATURES Location/Qualifiers

[illegible]

```
Qy   [illegible] Val[illegible]
Db   [illegible]GATA[illegible]4

RESULT 6
AX4[illegible]77[illegible]
LOCUS       AX4[illegible]   [illegible]4 bp    DNA    linear   PAT [illegible]
DEFINITION  Sequence [illegible] from Patent W[illegible]
ACCESSION   AX4[illegible]
VERSION     AX4[illegible]  GI:214[illegible]979
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tang,Y.T., Liu,C., Zhou,P., Asundi,V., Zhang,J., Zhang,Q.A., [illegible]
            Xue,A.J., Yang,Y., Wehrman,T. and [illegible]
  TITLE     Novel nucleic acids and polypeptides
  JOURNAL   Patent: W[illegible] A [illegible] MAR [illegible]
            HYSEQ, INC. (US)
FEATURES             Location/Qualifiers
     source          1..[illegible]
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             234..[illegible]2
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD34877.1"
                     /db_xref="GI:214[illegible]980"
                     /translation="MAAAME[illegible]
                     LKKLLAD[illegible]KYH[illegible]YMMAKAP[illegible]AMS[illegible]
                     PE[illegible]
                     IFLFQAGS[illegible]YHVK[illegible]PLFPNLVF[illegible]PNIFM[illegible]KD[illegible]WIA[illegible]
                     D[illegible]WISNIV[illegible]YVHNL[illegible]AVA[illegible]
                     PS[illegible]GK[illegible]LSEVE[illegible]ISPM[illegible]
                     DA[illegible]
                     PELFIPV[illegible]VMER[illegible]VPL[illegible]WINHH[illegible]
                     IIAS[illegible]CK[illegible]LYK[illegible]
                     HGSN[illegible]QV[illegible]VR[illegible]LVY[illegible]SYVN[illegible]
                     F[illegible]SKYSN[illegible]KNPH[illegible]VSLYKLSSP[illegible]KIKEEWA[illegible]LLDSA[illegible]
                     I[illegible]GF[illegible]LY[illegible]MIYKPH[illegible]GKKYP[illegible]VATAGAPV[illegible]LW[illegible]
                     P[illegible]N[illegible]SVAM[illegible]AEKE[illegible]NVHEA[illegible]
                     [illegible]AALKVI"
BASE COUNT    [illegible] a   592 c   [illegible] g   [illegible] t
ORIGIN

Alignment Scores:
Pred. No.:              [illegible]          Length:        [illegible]
Score:                  4118.[illegible]     Matches:       [illegible]
Percent Similarity:     [illegible]          Conservative:  [illegible]
Best Local Similarity:  [illegible]          Mismatches:    [illegible]
Query Match:            [illegible]          Indels:        [illegible]
DB:                     [illegible]          Gaps:          [illegible]

US-10-070-464-1 ([illegible]) x AX4[illegible] ([illegible])

Qy   [illegible] MetAlaAlaAlaMet[illegible]Asp[illegible]
Db   2[illegible] A[illegible]
Qy   [illegible] Asn[illegible]IleSer[illegible]AspAl[illegible]Asp[illegible]
Db   2[illegible]
Qy   [illegible] Ser[illegible]Lys[illegible]AlaAsp[illegible]Met[illegible]
Db   [illegible]
Qy   [illegible] Ala[illegible]Asp[illegible]Met[illegible]Val[illegible]Asp[illegible]
Db   41[illegible]
```

[illegible]

```
Qy    [illegible] ValIle [illegible]
Db    [illegible]AIA [illegible]

RESULT [illegible]
[illegible]
LOCUS       [illegible]          [illegible] bp    mRNA    linear   PRI [illegible] MAY [illegible]
DEFINITION  Homo sapiens, [illegible] mRNA, [illegible]
ACCESSION   [illegible]
VERSION     [illegible]
KEYWORDS    MGC.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   [illegible] (bases 1 to [illegible])
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (14 MAY 2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room [illegible], Bethesda, MD [illegible]
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: [illegible]
            Tissue Procurement: [illegible]
            cDNA Library Preparation: [illegible] (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: [illegible]
            http://www.[illegible]
            contact: [illegible]
            Anup Madan, Jessica [illegible], Mark [illegible]
            Madan, Stephanie Rodrigues, Amy [illegible] and Michelle White.

            [illegible]
            [illegible] I.M.A.G.E. [illegible]
            Series: IRAK [illegible]
```

```
Alignment Scores:
Pred. No.:    [illegible]
Score:        [illegible]
Percent Similarity: [illegible]
Best Local Similarity: [illegible]
Query Match:  [illegible]
DB:           [illegible]
```

[illegible]

```
SOURCE      [illegible]
  ORGANISM  [illegible]
REFERENCE   [illegible]
  AUTHORS   [illegible] V.S., [illegible] A.F., [illegible]
            [illegible] DeLeppaine,A.M., [illegible] M.F., Nguyen, [illegible]
            [illegible] F.A., Walker,N.K., Yao,M.[illegible], [illegible]
            [illegible] Y.[illegible], Walsh,F.[illegible], Azimzai,Y., Ramkumar,[illegible], Xu,Y. [illegible]
  JOURNAL   Patent: W[illegible] A[illegible] 27-DEC-2001;
            [illegible] Inc. (US)
FEATURES    [illegible]
BASE COUNT  [illegible]
ORIGIN
```

[illegible]

```
AX[illegible]
LOCUS       AX354[illegible]   [illegible] bp   [illegible]NA   linear   [illegible]
DEFINITION  Sequence [illegible] from Patent W[illegible]
ACCESSION   AX354[illegible]
VERSION     AX354[illegible]
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; [illegible]
REFERENCE   1
  AUTHORS   Meyers,R.A. and Williamson,M.
  TITLE     21953, a human prolyl oligopeptidase family member and uses [illegible]
  JOURNAL   Patent: WO [illegible]17917-A [illegible]
            Millennium Pharmaceuticals, Inc. [illegible]
FEATURES             Location/Qualifiers
     source          1..2[illegible]
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT  [illegible]
ORIGIN

Alignment Scores:
Pred. No.:              0              Length:        2[illegible]
Score:                  47[illegible]  Matches:       [illegible]
Percent Similarity:     [illegible]    Conservative:  [illegible]
Best Local Similarity:  [illegible]    Mismatches:    [illegible]
Query Match:            [illegible]    Indels:        [illegible]
DB:                     6              Gaps:          [illegible]

US-10-070-464-1 ([illegible]) x AX354[illegible] (1-2[illegible])

Qy          1 MetAlaAlaAlaMet[illegible]
Db          1 ATG[illegible]AAATGGGA[illegible]
[illegible]
```

protein - nucleic search, using frame plus p2n model

SUMMARIES

ALIGNMENTS

[illegible]

XX
AC [illegible] (first entry)
XX
KW Human; serine protease; dipeptidyl peptidase IV related protein; [illegible]
KW [illegible] infection; human immunodeficiency virus; HIV 1; HIV 2; [illegible]
KW [illegible] infertility; [illegible]
KW heart failure; hypertension; urinary retention; [illegible]
KW [illegible]
KW [illegible] reproductive disorder; inflammatory disorder;
KW metabolic disorder.
XX
[illegible]
XX
PN WO[illegible]-A1.
XX
[illegible]
XX
PA [illegible] NT BV.
XX
[illegible]
XX
DR [illegible]
XX
PT New dipeptidyl peptidase IV related protein [illegible]
PT [illegible]
PT [illegible]
XX
[illegible]
XX
CC The present invention relates to [illegible] serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV) related
CC proteins (DPRP). The dipeptidyl peptidase IV related proteins (DPRP)
CC [illegible]
CC [illegible] caused by human immunodeficiency virus (HIV [illegible])
CC [illegible] diabetes [illegible]
CC [illegible]
CC [illegible]
CC [illegible]
CC [illegible] disorders [illegible]
CC [illegible]
CC [illegible]
CC [illegible]
XX
SQ Sequence [illegible] AA;

[illegible]

[illegible]

[illegible]

RESULT 14
[illegible]
XX
[illegible]
XX
DT 12-APR-2002 (first entry)
XX
[illegible] splice variant #1.
XX
KW Human; serine protease; dipeptidyl peptidase IV related protein; [illegible]
KW [illegible]; infection; human immunodeficiency virus; HIV-1; HIV-2; [illegible]
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; [illegible]
KW heart failure; hypertension; urinary [illegible]; osteoporosis; [illegible]
KW [illegible]; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder.
XX
OS Homo sapiens.
XX
PN [illegible]
XX
PD [illegible]-APR-2002.
XX
[illegible]
XX
[illegible]
XX
PA (FERR-) FERRING BV.
XX
PI [illegible] Akinsanya K, [illegible]
XX
DR WPI; 2002-441178/47.
DR N-PSDB; ABK[illegible].
XX

[illegible]

XX
[illegible]
XX
[illegible]
[illegible]
XX
SQ Sequence [illegible] AA;

[illegible]

```
CC   [illegible] (DPRP) [illegible] dipeptidyl peptidase IV [illegible]
CC   and [illegible] encoding them are useful for treating infections
CC   such as fungal, bacterial, protozoan and viral infections, particularly
CC   infections caused by human immunodeficiency virus (HIV 1 or HIV 2),
CC   pain, diabetes, precocious puberty, infertility, obesity, anorexia,
CC   bulimia, Parkinson's disease, acute heart failure, hypotension,
CC   hypertension, urinary retention, osteoporosis, angina [illegible]
CC   [illegible] ulcers, asthma, allergies, [illegible]
CC   [illegible] and neurological disorders [illegible]
CC   [illegible] dyskinesias. These [illegible]
CC   therapeutic agents for the treatment of reproductive, inflammatory and
CC   metabolic disorders. [illegible] represent human DPRP proteins.
XX
SQ   Sequence [illegible] AA;

[illegible]
[illegible]
[illegible]

Qy   [illegible] MAAAMETEQLGVEIFETADCEENIEQPDEKENPPYVFEYLWKQKKELALEKYEFYMM [illegible]
Db   [illegible] MAAAMETEQLGVEIFETADCEENIEQPDEKENPPYVFEYLWKQKKELALEKYEFYMM [illegible]

Qy   [illegible] MEVKKNEEPEHEDEKEVYYLAMI [illegible] INKAAVMEWEI [illegible]
Db   [illegible] MEVKKNEEPEHEDEKEVYYLAMI [illegible] INKAAVMEWEI [illegible]

Qy   [illegible] ATLLYMMYSEEEEEKEKEIIVTIA [illegible] YHVKE [illegible]
Db   [illegible] ATLLYMMYSEEEEEKEKEIIVTIASYYHQESIEEPQASEIYHVKESHPQE [illegible]

Qy   [illegible] LEEENEVELI [illegible] NIEMEPKE [illegible] AEEWIAE [illegible] IVBNE [illegible]
Db   [illegible] LEEENEVELI [illegible] NIEMEPKE [illegible] AEEWIAE [illegible] IVBNE [illegible]

Qy   [illegible] ANMEEAEGAVAIEVL [illegible] YWW [illegible] PKAEE [illegible] LYEENEEEVEEIHV [illegible]
Db   [illegible] ANMEEAEGAVAIEVLQEEEDKYSYWWTPKAEEIPSMEEIRILYEENEEEVEEIHV [illegible]

Qy   [illegible] MEEEEAECEPYEKE [illegible] ANEKVIEKMEEMEEAE [illegible] VIEKPE [illegible]
Db   [illegible]

Qy   [illegible] KAEAWEE [illegible] KYAWEE [illegible] VMEEQEEEEVEE [illegible]
Db   [illegible] KAEAWEE [illegible] KYAWEEE [illegible] VMEEQEEEEVEEIVEEE [illegible]

Qy   [illegible] WINEE [illegible]
Db   [illegible] WINEE [illegible]

Qy   [illegible]
Db   [illegible]

Qy   [illegible] SVNE [illegible]
Db   [illegible] SVNE [illegible]

Qy   [illegible] WA [illegible]

Db   [illegible] WAIE [illegible]

[illegible]
[illegible]
[illegible] AA
XX
[illegible]
XX
[illegible] (first entry)
XX
[illegible]
XX
KW   Human [illegible] dipeptidyl peptidase IV related [illegible]
KW   DPPIV [illegible] human immunodeficiency virus; HIV 1; HIV 2; [illegible]
KW   diabetes; [illegible] obesity; anorexia; Parkinson's disease; [illegible]
KW   heart failure; hypertension; urinary retention; osteoporosis; [illegible]
KW   ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW   dyskinesia; reproductive disorder; inflammatory disorder;
KW   metabolic disorder.
XX
OS   Homo sapiens.
XX
FN   W [illegible]
XX
PD   [illegible]
XX
PF   [illegible]
XX
PR   [illegible]
XX
PA   [illegible]
XX
PI   [illegible] Akiba [illegible]
XX
DR   WPI; [illegible]
DR   N-PSDB; [illegible]
XX
PT   New dipeptidyl peptidase IV related proteins and nucleic acids [illegible]
PT   the proteins [illegible] treating [illegible] fungal, bacterial, protozoan and
PT   viral infections, cancers, allergies, neurological disorders, [illegible]
PT
XX
PS   Claim [illegible]; English.
XX
CC   The present invention relates to the isolation of novel human [illegible]
CC   proteins referred to as dipeptidyl peptidase IV (DPPIV)-related
CC   proteins (DPRP). The dipeptidyl peptidase IV related proteins (DPRP)
CC   and nucleic acids encoding them are useful for treating infections
CC   such as fungal, bacterial, protozoan and viral infections, particularly
CC   infections caused by human immunodeficiency virus (HIV 1 or HIV 2),
CC   pain, diabetes, precocious puberty, infertility, obesity, anorexia,
CC   bulimia, Parkinson's disease, acute heart failure, hypotension,
CC   hypertension, urinary retention, osteoporosis, angina pectoris,
CC   [illegible] ulcers, asthma, allergies, [illegible]
CC   psychotic and neurological disorders (e.g. anxiety, dementia [illegible]
CC   schizophrenia, [illegible] dyskinesias. These may also be used in [illegible]
CC   therapeutic agents for the treatment of reproductive, inflammatory and
CC   metabolic disorders. [illegible] represent human DPRP proteins.
XX
SQ   Sequence [illegible] AA;

[illegible]
[illegible]
Matches [illegible]

Qy   [illegible] EKYHEYHMAKAEHEEMEVKENDPEGHEDEEEYYLAM [illegible]
Db   [illegible] VNKAEHE [illegible] IVKEEE [illegible] YYL [illegible]

Qy   [illegible] ATLLYMMYSEEEEEKEKEIIV [illegible]
Db   [illegible] VYSEEEEEEEKEEDVE [illegible]

Qy   [illegible] HVKE [illegible] LEEENEVELISENIEMEPKE [illegible] AEEWIAE [illegible]
Db   [illegible] PKNEMV [illegible] EKE [illegible] EMEPKE [illegible] AEEAE [illegible]

Qy   [illegible] VVBNE [illegible] ANMEEAEGAVAIEVL [illegible] YWW [illegible]
Db   [illegible] SNV [illegible] AEGAVAIEVLEEEDEE [illegible] YWW [illegible]

Qy   [illegible] QEEEVEEIHVEEMEEEEEEAEQEPYEKEIGANEKVIEKMEEIM [illegible]
Db   [illegible] EVEVIHVEEAEEEKEE [illegible] KNEKIA [illegible]
```

PD [illegible]
XX
PF [illegible]
XX
PR 2? JUL [illegible]; [illegible]JP-[illegible].
PR ?-AUG [illegible]; [illegible]JP-[illegible].
PR [illegible]-JAN [illegible]; [illegible]JP-[illegible].
PR [illegible]-MAY [illegible]; [illegible]JP-[illegible].
PR [illegible]-JUN-[illegible]; [illegible]JP-[illegible].
XX
PA (HELI) HELIX RES INST.
XX
PI [illegible] Isogai T, Nishikawa T, Hayashi K, [illegible], Yamamoto
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, [illegible]saki [illegible]
XX
DR WPI; [illegible]/[illegible]4.
XX
PT Primer sets for synthesizing [illegible], particularly [illegible]
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim [illegible]; SEQ ID [illegible]; [illegible]pp; [illegible] English.
XX
CC The present invention describes primer sets for synthesising the
CC full-length cDNAs defined in the specification, where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 562 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy or
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH[illegible] to AAH[illegible] and
CC AAH[illegible] to AAH[illegible]742 represent human cDNA sequences; AAH[illegible] to
CC AAH[illegible] represent human amino acid sequences; and AAH[illegible] to AAH[illegible]
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence [illegible] AA;

Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy 2[illegible] MEPKI[illegible]WIAHHSNI[illegible]WISNIVIPEERR[illegible]VHN[illegible]ANME[illegible]DARSA[illegible]VAIEV[illegible]

Db 1 ME[illegible]KI[illegible]WIAHHSNI[illegible]WISN[illegible]VIPEE[illegible]V[illegible]NE[illegible]ANME[illegible]DARSA[illegible]VAIEV[illegible]

Qy 2[illegible] EPDRY[illegible]YWA[illegible]KAEI[illegible]KILE[illegible]YEEN[illegible]HVL[illegible]HE[illegible]RKA[illegible]

Db [illegible] EE[illegible]RYSIYWA[illegible]KAEI[illegible]PS[illegible]KIL[illegible]YEEN[illegible]AVI[illegible]EM[illegible]EEA[illegible]SEFY[illegible]

Qy [illegible] IANPKVI[illegible]KMSLIMI[illegible]A[illegible]VI[illegible]KEL[illegible]VEY[illegible]A[illegible]W[illegible]KYAWS[illegible]

Db 1[illegible] IANEKVI[illegible]KMS[illegible]IMI[illegible]A[illegible]VI[illegible]KEL[illegible]VEY[illegible]A[illegible]W[illegible]KYAWS[illegible]

Qy [illegible] [illegible]S[illegible]IE[illegible]VI[illegible]SP[illegible]VE[illegible]VM[illegible]SV[illegible]VIP[illegible]YE[illegible]IWINIH[illegible]

Db 1[illegible] [illegible]SQIEF[illegible]IVEISPELEP[illegible]VM[illegible]SV[illegible]V[illegible]YE[illegible]IWINIHD[illegible]

Qy [illegible] [illegible]SHEE[illegible]ASE[illegible]K[illegible]HY[illegible]KE[illegible]AESE[illegible]KEEIA[illegible]

Db 2[illegible] EP[illegible]SHEEEIE[illegible]EASE[illegible]KI[illegible]HLYKIISIEKESKY[illegible]AESE[illegible]K[illegible]KEEIA[illegible]

Qy [illegible]

Db [illegible]

Qy [illegible] KY[illegible]KN[illegible]

Db [illegible] Y[illegible]

Qy [illegible]

Db [illegible]

Qy [illegible]

Db 4[illegible] YVVVVI[illegible]NE[illegible]KEE[illegible]

Qy [illegible]

Db [illegible]

Qy [illegible]

Db [illegible]

Qy [illegible]

Db [illegible]

RESULT [illegible]
AA[illegible]
ID AA[illegible]
XX
AC AA[illegible]
XX
DT [illegible]
XX
DE Human [illegible]
XX
KW [illegible]
KW [illegible]
KW [illegible]
KW [illegible]
KW [illegible]
KW [illegible]
KW [illegible]
XX
OS Homo sapiens.
XX
PN W[illegible]
XX
PD [illegible]
XX
PF [illegible]
XX
PR [illegible]
XX
PA (HEFE) [illegible]
XX
PI [illegible]
XX
DR WPI; [illegible]
DR N-PSDB; AAK[illegible]
XX
PT [illegible]
PT [illegible]
PT [illegible]
PT
XX
PS [illegible]
XX
CC The present [illegible]
CC [illegible]

[illegible]

XX

[illegible]

Human [illegible] variant #1.

[illegible]

XX

[illegible]

Qy 1 MAAAM[illegible]

[illegible]

Qy 361 YIAKA[illegible]

[illegible]

AAB[illegible]

XX

[illegible]

XX

```
CC   pain, diabetes, [illegible] obesity, [illegible]
CC   bulimia, Parkinson's disease, [illegible] heart failure, hypotension,
CC   hypertension, urinary retention, osteoporosis, angina pectoris,
CC   stroke, ulcers, asthma, allergies, [illegible]
CC   psychotic and neurological disorders [illegible]
CC   [illegible] and dyskinesias. These [illegible] used [illegible]
CC   therapeutic agents for the treatment [illegible]
CC   metabolic disorders. AAW[illegible] represent human [illegible] proteins
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy          1 MAAAME[illegible]VEIFE[illegible]EENIES[illegible]EKLEE[illegible]YV[illegible]LKK[illegible]A[illegible]KYH[illegible]
Db          1 MAAAME[illegible]VEIF[illegible]EENIES[illegible]KLE[illegible]YV[illegible]LKK[illegible]A[illegible]EKYH[illegible]YMM

Qy         61 AKA[illegible]EM[illegible]VKEN[illegible]YYIAMS[illegible]LNEEN[illegible]K[illegible]NRAAV[illegible]WKE
Db         61 AKA[illegible]EM[illegible]VKEN[illegible]YYIAMS[illegible]ENEEN[illegible]K[illegible]NRAAV[illegible]WKE

Qy        121 [illegible]A[illegible]MYSREE[illegible]EKK[illegible]VG[illegible]ASY[illegible]AS[illegible]YH[illegible]K[illegible]
Db        121 [illegible]A[illegible]GMYSREEFLLE[illegible]KK[illegible]VG[illegible]ASY[illegible]AS[illegible]YHVK[illegible]

Qy        181 EIQQPLRPNLVETSCPNIRM[illegible]PKICPA[illegible]PEW[illegible]AE[illegible]N[illegible]IW[illegible]NIVIRE[illegible]IYVHN[illegible]
Db        181 EIQ[illegible]NLVETSCPNIEM[illegible]PKI[illegible]PA[illegible]EW[illegible]AE[illegible]N[illegible]IW[illegible]NIVI[illegible]IYVEN[illegible]

Qy        241 ANMEE[illegible]A[illegible]VAIFVI[illegible]YWW[illegible]PKAE[illegible]K[illegible]YEENG[illegible]
Db        241 ANMEE[illegible]A[illegible]VAIFVI[illegible]EEF[illegible]FYS[illegible]YWW[illegible]PKAF[illegible]K[illegible]YEENG[illegible]

Qy        301 [illegible]SPMLE[illegible]RAL SERYPKI[illegible]ANEKVIEKMSEIM[illegible]AE[illegible]VI[illegible]KELI[illegible]LLE[illegible]VE
Db        [illegible] [illegible]EMSE[illegible]IE[illegible]SERY[illegible]ANEKVIEKMSEIM[illegible]AE[illegible]VIDKELIQPE[illegible]LLEP[illegible]VE

Qy        361 YIAEA[illegible]WIEE[illegible]KYAWSILLE[illegible]IV[illegible]IS[illegible]VM[illegible]
Db        361 YIAEA[illegible]WIEE[illegible]KYAWSILLE[illegible]IV[illegible]IS[illegible]VM[illegible]

Qy        421 [illegible]YEL[illegible]IWINIHD[illegible]EHVE[illegible]SHE[illegible]ASE[illegible]YK[illegible]SILKE[illegible]KYKE[illegible]
Db        421 [illegible]YLEL[illegible]IWINIHDIFHVIE[illegible]SHEE[illegible]EIEASE[illegible]K[illegible]YK[illegible]SILKE[illegible]KYKES[illegible]

Qy        481 [illegible]KEE[illegible]AII[illegible]WEV[illegible]SN[illegible]V[illegible]KE[illegible]YVVS
Db        481 [illegible]A[illegible]KEE[illegible]AII[illegible]WEV[illegible]SN[illegible]V[illegible]KE[illegible]YVVS

Qy        541 YVNP[illegible]VI[illegible]YSHS[illegible]SKY[illegible]N[illegible]YKISSPEE[illegible]KIKE[illegible]
Db        541 YVNP[illegible]VI[illegible]GYSHS[illegible]SKYSN[illegible]KN[illegible]V[illegible]LYKLSSPEE[illegible]KIKE[illegible]

Qy        601 WA[illegible]SA[illegible]PDYIPPE[illegible]SPES[illegible]YCMLYK[illegible]KKYFIVLE[illegible] 655
Db        601 WA[illegible]SA[illegible]PDYIPPE[illegible]SEESI[illegible]YCMLYK[illegible]KKYFIVLE[illegible]

RESULT [illegible]
AB[illegible]
ID   AB[illegible] standard; [illegible] AA.
XX
AC   AB[illegible];
XX
DT   [illegible] (first entry)
XX
DE   Human [illegible] splice variant [illegible]
XX
KW   Human; [illegible] protease; [illegible] peptide; [illegible]
KW   [illegible] human [illegible]
KW   diabetes; infertility; obesity; [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
XX
OS   [illegible]
XX
PN   W[illegible]
XX
PD   [illegible]
XX
PF   [illegible]
XX
PA   [illegible]
XX
PI   [illegible]
XX
DR   W[illegible]
DR   N-PSDB; [illegible]
XX
PT   New [illegible]
PT   [illegible]
PT   [illegible]
PT   [illegible]
XX
PS   [illegible]
XX
CC   The present [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   pain, diabetes, [illegible]
CC   bulimia, Parkinson's disease, [illegible]
CC   hypertension, urinary retention, [illegible]
CC   stroke, [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   metabolic [illegible]
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy          1 MAAAME[illegible]VE[illegible]
Db          1 MAAAME[illegible]VE[illegible]YV[illegible]LKK[illegible]A[illegible]

Qy         61 AKA[illegible]EMVKEN[illegible]NRAAV[illegible]
Db         61 AKA[illegible]EMVKEN[illegible]NRAAV[illegible]

Qy        121 [illegible]
Db        121 [illegible]A[illegible]

Qy        181 [illegible]
Db        181 EIQ[illegible]PNLVE[illegible]

Qy        241 [illegible]
Db        241 [illegible]ANM[illegible]

Qy        [illegible]
Db        [illegible]
```

```
XX
XX
XX
XX
XX
XX
PT   [illegible]
XX
DE   An isolated [illegible] for treating [illegible] associated with the
DE   [illegible] polypeptide such as cancer and multiple [illegible]
XX
PS   [illegible]
XX
CC   The present [illegible] provides [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible] treatment, to regulate the [illegible] system [illegible]
CC   [illegible] to regulate [illegible]
CC   [illegible] to regulate haemostasis [illegible] thrombolysis [illegible]
CC   stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC   [illegible] rheumatoid arthritis, and to treat nervous system disorders [illegible]
CC   Parkinson's disease. The present sequence is a protein of the invention.
XX
SQ   Sequence [illegible] AA;

Qy        1 MAAAMETE[illegible]
Db        1 MAAAMETE[illegible]

Qy       [illegible] AKAPHEEMEVKRNPEPEHHDEIYYLAMSPENRENILEYSEIPKEINRAAVIMLSWKEIE [illegible]
Db       [illegible] AKAPHE[illegible]

Qy       [illegible] ANMEEPARAAVATEV[illegible]EYSYWW[illegible]PKAP[illegible]
Db       [illegible] ANMEEPARAAVATEV[illegible]

Qy       [illegible] VATAGAPVITWIFYEI[illegible]
Db       [illegible] VATAGAPVITWIFYEI[illegible]

Qy       [illegible] NVHFAH[illegible]
Db       [illegible] NVHFAH[illegible]

Qy       [illegible] HYEIHDIRYL[illegible]PNI[illegible]AAIKVI [illegible]
Db       [illegible] HYEIHDIRYL[illegible]PNI[illegible]AAIKVI [illegible]


RESULT [illegible]
ABG[illegible]
ID   ABG[illegible] standard; protein; [illegible] AA.
XX
AC   ABG[illegible]
XX
DT   [illegible]
XX
DE   Human [illegible] variant #[illegible]
XX
KW   Human; [illegible] dipeptidyl peptidase IV related protein; [illegible]
KW   DPPIV; [illegible] human immunodeficiency virus; HIV [illegible] HIV [illegible]
KW   diabetes; [illegible] obesity; [illegible] Parkinson's disease; [illegible]
KW   heart [illegible] hypertension; asthma; retention; osteoporosis; cancer; [illegible]
KW   ulcer; [illegible] psychotic disorder; neurological disorder;
KW   dyskinesia; [illegible] disorder; inflammatory disorder;
KW   metabolic [illegible]
XX
OS   Homo sapiens.
XX
PN   W[illegible]
XX
XX
XX
XX
PA   (PPEE[illegible]
XX
PI   [illegible]
XX
DR   [illegible]
XX
PT   New [illegible] dipeptidyl peptidase IV related proteins and [illegible]
PT   [illegible]
PT   [illegible] neurological [illegible]
PT
XX
PS   [illegible] English.
XX
CC   The present invention relates to the [illegible] of novel human [illegible]
CC   [illegible] dipeptidyl peptidase IV (DPPIV) related
CC   [illegible] dipeptidyl peptidase IV related [illegible]
CC   [illegible] useful for treating [illegible]
CC   [illegible] in particular
CC   [illegible] immunodeficiency virus (HIV [illegible]
```

```
Db      [illegible] ASRYIE [illegible] WSY [illegible] ISLMAIM [illegible] YVSIA [illegible] V [illegible] Y [illegible]

Qy      [illegible] NBFQNEL YYLTSVAMQAEKFFSEENFELLHHCHEENVHHAHISTL [illegible] LVEAGKEYI [illegible]

Db      [illegible] NHFQNEL YYLTSVAMQAEKFFSEENFELLHHCHEENVHHAHISTLLSEEVEAGKE [illegible]

Qy      [illegible] LIYFQEHHSIEVFESEHHYEHEHLHYLQEN [illegible] SEIAAL [illegible] V [illegible]

Db      [illegible] LIYFQEHHSIEVFESEHHYEHEHLHYLQENQSEIAALKV [illegible]


RESULT [illegible]
ABB67[illegible]
ID   ABB67[illegible] standard; Protein; [illegible] AA.
XX
AC   ABB67[illegible];
XX
DT   [illegible] (first entry)
XX
DE   Novel human protein SEQ ID NO [illegible].
XX
KW   Human; antianaemic; [illegible] antiinflammatory; [illegible]
KW   antiinfertility; cerebroprotective; cytostatic; [illegible] rheumatic; [illegible]
KW   neuroprotective; antiparkinsonian; [illegible] therapy; [illegible]
KW   expressed sequence tag.
XX
OS   Homo sapiens.
XX
PN   WO[illegible].
XX
PD   [illegible]
XX
PF   [illegible] 2001; 2001WO-US26[illegible].
XX
PR   [illegible] US-[illegible].
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT, [illegible], Zhou P, Asundi V, [illegible], Zhao QA, [illegible]
PI   Xue AJ, Yang Y, Wehrman T, [illegible]
XX
DR   WPI; 2002-[illegible]/[illegible].
DR   N-PSDB; ABN[illegible].
XX
PT   An isolated [illegible] for treating diseases associated with [illegible]
PT   encoded polypeptide such as cancer and multiple sclerosis.
XX
PS   Example [illegible]; SEQ ID NO [illegible]; [illegible] pp; English.
XX
CC   The present invention provides the protein and nucleic acid sequences [illegible]
CC   [illegible] human [illegible]. These were isolated [illegible] expressed [illegible]
CC   (ESTs). They can be used to stimulate [illegible] growth, to regulate
CC   [illegible] to treat [illegible]
CC   [illegible]
CC   [illegible] tissues, to regulate [illegible] to treat
CC   infertility, to regulate haemostasis or thrombolysis, [illegible] to treat
CC   stroke and cancer, [illegible] to treat [illegible]
CC   e.g. rheumatoid arthritis, and to treat [illegible] nervous system disorders [illegible]
CC   Parkinson's disease. The present sequence is a protein of the invention
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
  Best Local Similarity [illegible]; Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy      1 MAAAMETCQGVEIHIAE [illegible] ENIE [illegible] LVEAGKEIAIKEYKHYM [illegible]

Db      1 MAAAMETCQGVEIHIAEMLNNIE [illegible] LVEAGKEIAIKEYKHYM [illegible]

Qy      [illegible] AFARDIMSVKENIE [illegible] LAMS [illegible] N [illegible] NAAAINHSW [illegible]
```

```
Db      [illegible] AFA [illegible] NSVKENI [illegible]

Qy      [illegible]

Db      [illegible]

Qy      [illegible] NVEI [illegible]

Db      [illegible]

Qy      [illegible] ANM [illegible] VA [illegible]

Db      [illegible] ANM [illegible] VA [illegible]

Qy      [illegible]

Db      [illegible]

Qy      [illegible]

Db      [illegible]

Qy      [illegible]

Db      [illegible]

Qy      [illegible]

Db      [illegible]

Qy      [illegible] VAN [illegible]

Db      [illegible] VAN [illegible]

Qy      [illegible] WA [illegible]

Db      [illegible] WAL [illegible] EKY [illegible]

Qy      [illegible] VAN [illegible]

Db      [illegible]

Qy      [illegible]

Db      [illegible]

Qy      [illegible] NHFQNEL YYLTSVAM [illegible] NVHHAHISTL [illegible]

Db      [illegible]

Qy      [illegible]

Db      [illegible]


RESULT [illegible]
ABB[illegible]
ID   ABB[illegible]
XX
AC   ABB[illegible]
XX
DT   [illegible]
XX
DE   Novel [illegible] NO [illegible]
XX
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
XX
OS   [illegible]
XX
PN   W[illegible]
```

[illegible]

XX
DE Amino acid sequence of 21[illegible] human prolyl [illegible]peptidase.
XX
KW [illegible]
KW [illegible] disorder; inflammatory disease; [illegible]
KW [illegible] disorder; vascular disorder; prostate disorder; [illegible]
KW [illegible]
KW [illegible] asthma; [illegible] multiple sclerosis; [illegible]
KW [illegible] disease; neurological disorder; demyelinating disease.
XX
[illegible]
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
[illegible]
XX

[illegible]

CC [illegible] and
CC metabolic [illegible] uses include gene therapy, expression [illegible]
CC [illegible] modulators. It is used [illegible] identify [illegible]
CC compounds [illegible] and can be used in preventing/treating
CC [illegible] cellular proliferation or differentiation disorders.
CC The 21[illegible] molecules can act as novel diagnostic targets and therapeutic
CC agents [illegible] disorders associated with the aberrant [illegible]
CC [illegible] peptide hormones e.g. disorders associated with [illegible]
CC differentiation and proliferation such as cancer, immune function,
CC reproduction, [illegible]logical and cardiovascular function. The 21[illegible]
CC molecules are [illegible] useful for treating and preventing cellular
CC proliferation and differentiative disorders, haematopoietic [illegible]
CC disorders, immune disorders such as autoimmune diseases, diabetes
CC mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,
CC [illegible] demyelinating diseases, vascular disorders and
CC metabolic [illegible] disorders. This sequence represents the amino
CC acid sequence [illegible] human prolyl [illegible]peptidase.
XX
SQ Sequence [illegible]

[illegible]

Qy [illegible] MAAA[illegible] [illegible]

[illegible]

```
Db    241 ANMELIA[illegible]VATIV[illegible]WW[illegible]TKAR[illegible]

Qy    361 ISEMLEIE[illegible]AISEKYPKIGIANPKVIIKMSLIMIIAL[illegible]V[illegible]
Db    361 ISEMLEI[illegible]AISEEYIKIGIAN[illegible]KV[illegible]KM[illegible]MIIAL[illegible]

Qy    361 YIALA[illegible]W[illegible]KYAW[illegible]IV[illegible]VM[illegible]
Db    361 YIAKA[illegible]W[illegible]KYAW[illegible]IV[illegible]VM[illegible]

Qy    421 [illegible]YEEI[illegible]IWINIHD[illegible]HVI[illegible]SHI[illegible]ASD[illegible]YK[illegible]SILKE[illegible]SKYK[illegible]
Db    421 [illegible]YEEI[illegible]IWINIHD[illegible]HVI[illegible]SHI[illegible]ASD[illegible]YK[illegible]SILKE[illegible]SKYK[illegible]

Qy    481 [illegible]PAPSIEK[illegible]IKEEIATISGEWEVLGR[illegible]SNI[illegible]V[illegible]
Db    481 [illegible]AESIEK[illegible]IKEEIAT[illegible]EWEV[illegible]GR[illegible]

Qy    541 YVNP[illegible]EV[illegible]ILRGYSHS[illegible]HC[illegible]FISKYSN[illegible]KN[illegible]VS[illegible]YK[illegible]SHE[illegible]KIFL[illegible]
Db    541 YVNP[illegible]EV[illegible]LDRGYSHS[illegible]HC[illegible]FISKYSN[illegible]KN[illegible]VS[illegible]YK[illegible]SPE[illegible]KIKE[illegible]

Qy    601 WAIILDSA[illegible]LPDYIPPEIFSPESI[illegible]GFILYGMLYKEH[illegible]PGKKYP[illegible]VLI[illegible]V[illegible]
Db    601 WAIILDSA[illegible]LPDYIPPEIFSPESI[illegible]GFILYGMLYKEH[illegible]PGKKYP[illegible]VI[illegible]V[illegible]

Qy    661 VNNPFK[illegible]VKYTRLNILAS[illegible]YVVVV[illegible]NR[illegible]HR[illegible]AEKYKM[illegible]VE[illegible]
Db    661 VNNPFK[illegible]KYTRLNILA[illegible]YVVVV[illegible]LNR[illegible]

Qy    721 LASRYI[illegible]EV[illegible]MALM[illegible]VA[illegible]IWIPY[illegible]YIT[illegible]
Db    721 LASRYI[illegible]EV[illegible]MALM[illegible]VA[illegible]

Qy    781 [illegible]NE[illegible]YYLGSVAMQAEKI[illegible]NRI[illegible]NV[illegible]AH[illegible]VPA[illegible]
Db    781 [illegible]NE[illegible]YYLGSVAMQAEKI[illegible]NRL[illegible]NV[illegible]AH[illegible]VPA[illegible]

Qy    841 [illegible]YPQ[illegible]RVP[illegible]HY[illegible]N[illegible]AA[illegible]KV[illegible]
Db    841 [illegible]YPQ[illegible]RVP[illegible]HYL[illegible]EN[illegible]AA[illegible]KV[illegible]
```

```
RESULT 4
AA[illegible]
ID   AAG[illegible] standard; Protein; 88[illegible] AA.
XX
AC   AA[illegible];
XX
DT   [illegible] Apr 2002 (first entry)
XX
DE   Human protease PRTS [illegible] protein sequence.
XX
KW   Human protease PRTS; [illegible]
KW   cardiovascular; atherosclerosis; [illegible] dermatitis;
KW   inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
KW   cell proliferative disorder; [illegible] epilepsy;
KW   [illegible] muscular dystrophy; epithelial [illegible]
KW   reproductive disorder; endometriosis.
XX
OS   Homo sapiens.
XX
PN   WO[illegible]
XX
PD   [illegible] 2001.
XX
PF   [illegible]
XX
PR   [illegible]
PR   [illegible]
PR   [illegible]
PR   [illegible]
```

```
[illegible]
XX
PA   [illegible]
XX
PI   [illegible]
XX
DR   [illegible]
XX
PT   [illegible]
XX
PS   [illegible]
XX
CC   The present [illegible]
CC   [illegible]
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy    [illegible]
Db    [illegible]
```

[illegible]

XX
[illegible]
XX
KW [illegible] dipeptidyl peptidase IV related [illegible]
KW [illegible] human immunodeficiency virus; HIV 1; HIV 2; [illegible]
KW diabetes; [illegible] obesity; [illegible] Parkinson's disease; [illegible]
KW heart failure; [illegible]
KW [illegible]
KW dyskinesia; [illegible] inflammatory disorder;
KW metabolic disorder.
XX
[illegible]
XX
XX
[illegible]
XX
[illegible] New dipeptidyl peptidase IV related proteins [illegible]
XX
[illegible]
XX
The present invention relates to the isolation of novel human serine proteases referred to as dipeptidyl peptidase IV (DPPIV) related proteins (DPRP). The dipeptidyl peptidase IV related proteins (DPRP) and [illegible] encoding them are useful for treating infections such as [illegible] protozoan and viral infections, particularly infections caused by human immunodeficiency virus (HIV 1 or HIV 2), pain, diabetes, precocious puberty, infertility, obesity, anorexia, [illegible] disease, acute heart failure, hypotension, hypertension, [illegible] asthma, [illegible] stroke, [illegible] migraine, [illegible] psychiatric [illegible] disorders (e.g. anxiety, dementia or schizophrenia) and dyskinesias. These may also be used in [illegible] therapeutic [illegible] for the treatment of reproductive, inflammatory and metabolic [illegible]. [illegible] represent human DPRP proteins.
XX
[illegible]

[illegible]

XX
PA (UNSY) UNIV SYDNEY.
XX
PI [illegible] MD;
XX
XX
PT New human dipeptidyl aminopeptidase [illegible], useful for [illegible]
PT substrates, identifying inhibitors of [illegible] catalytic activity which
PT have therapeutic uses, and for detecting activated [illegible] cells
XX
PS [illegible]
XX
CC This sequence represents human dipeptidyl aminopeptidase [illegible]
CC [illegible] has substrate specificity for H Ala Pro pNA, H Gly Pro pNA and
CC H Arg Pro pNA. Therefore, it is a prolyl oligopeptidase and a
CC dipeptidyl peptidase, because it is capable of hydrolysing the
CC peptide bond C-terminal to Pro in each of these compounds. [illegible]
CC is homologous with human DPP IV. [illegible] is useful for [illegible]
CC substrates, and for detecting an activated [illegible] cell, which involves
CC measuring the level of [illegible] gene expression [illegible]
CC of [illegible] expression is detected by detecting the amount of [illegible] RNA
CC in the cell. It is also useful for identifying a molecule capable
CC of inhibiting the cleavage of the substrate by [illegible]. Molecules
CC identified as inhibiting [illegible] catalytic activity may be used [illegible]
CC treating [illegible]
CC [illegible] insulin dependent diabetes [illegible] and other disorders
CC [illegible]
CC as immunosuppressants.
XX
SQ Sequence [illegible] AA;

Query Match [illegible] Score [illegible]
Best Local Similarity [illegible]
Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible]

Qy 1 MAAA[illegible]
Db 1 MAAAM[illegible]

Qy [illegible] AKAH[illegible]
Db [illegible] AKAH[illegible]

[illegible]

RESULT [illegible]
AAE[illegible]
XX
XX
XX
XX
KW [illegible]
XX
XX
XX
XX
PA (UNSY) UNIV SYDNEY.
XX
XX
XX
XX
XX
SQ Sequence [illegible] AA;

D-gene

SUMMARIES

ALIGNMENTS

```
RESULT [illegible]
AAH[illegible]
ID   AAH[illegible] standard; cDNA; [illegible] BP.
XX
AC   AAH[illegible];
XX
DT   [illegible] JUN [illegible] (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:[illegible].
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy;
XX
OS   Homo sapiens.
XX
PN   EP[illegible] A2.
XX
PD   [illegible] FEB [illegible].
XX
PF   28 JUN 2000; 2000EP-[illegible].
XX
PR   [illegible] JUL [illegible]; [illegible]JP-[illegible].
PR   [illegible] AUG [illegible]; [illegible]JP-[illegible].
PR   [illegible] JAN 2000; [illegible]JP-[illegible].
PR   [illegible] MAY 2000; 2000JP-[illegible].
PR   [illegible] JUN 2000; 2000JP-[illegible].
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J,
PI   Ishii S, Sugiyama T, Wakamatsu A, Nagai K, [illegible]
XX
DR   WPI; [illegible]/[illegible].
XX
PT   Primer sets for synthesizing polynucleotides, particularly the [illegible]
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 9; SEQ ID 12963; 2537pp; [illegible]; English.
XX
CC   The present invention describes primer sets for synthesising [illegible]
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo(dT) primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises [illegible]
CC   the 5' end nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5' end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3' end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5' end sequence/3' end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH[illegible] to AAH[illegible] and
CC   AAH[illegible] to AAH[illegible]742 represent human cDNA sequences; AAH[illegible] to
CC   AAH[illegible] represent human amino acid sequences and AAH[illegible] to AAH[illegible]
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence [illegible] BP; [illegible] A; [illegible] C; [illegible] G; [illegible] T; [illegible] other;

Alignment Scores:
Pred. No.:              [illegible]      Length:        [illegible]
Score:                  [illegible]      Matches:       [illegible]
Percent Similarity:     [illegible]      Conservative:  [illegible]
Best Local Similarity:  [illegible]      Mismatches:    [illegible]
Query Match:            [illegible]      Indels:        [illegible]
DB:                     22               Gaps:          [illegible]
```

```
Qy    [illegible]
Db    [illegible]

XX
AC    ABK[illegible];
XX
DT    [illegible] 2002 (first entry)
XX
DE    DNA encoding human [illegible] splice variant #[illegible].
XX
KW    Human; serine protease; [illegible] peptidase IV-related protein; [illegible]
KW    [illegible]; infection; human immunodeficiency virus; HIV [illegible] HIV [illegible]
KW    diabetes; infertility; obesity; anorexia; [illegible] disease; [illegible]
KW    heart failure; hypertension; urinary [illegible]
KW    ulcer; allergy; cancer; psychotic disorder; neurological disorder
KW    dyskinesia; [illegible] disorder; inflammatory [illegible];
KW    metabolic disorder; [illegible] ss.
XX
OS    Homo sapiens.
XX
PN    W[illegible] [illegible] A2.
XX
PD    [illegible]
XX
PF    [illegible] 2001; 2001WO-US31474.
XX
PR    [illegible]
XX
PA    [illegible]
XX
```

```
XX
XX
XX
CC    [illegible]
CC    protease [illegible]
CC    [illegible]
CC    such as [illegible]
CC    [illegible] HIV [illegible]
CC    [illegible] diabetes [illegible]
CC    [illegible]
CC    hypertension [illegible]
CC    stroke [illegible]
CC    psychotic [illegible]
CC    [illegible]
CC    therapeutic agents [illegible]
CC    [illegible]
XX

Alignment Scores:
[illegible]
Query Match: [illegible]
```

```
XX
XX
PD   [illegible] MAY [illegible]
XX
PF   [illegible] MAY [illegible]
PR   [illegible] MAY [illegible]
PR   [illegible] MAY [illegible]
PR   [illegible] MAY [illegible]
XX
PA   [illegible] INCYTE GENOMICS INC.
XX
PI   [illegible] AM, [illegible] Khan FA, [illegible] MB, [illegible]
XX
DR   WPI; [illegible]
XX
PT   Novel human proteases and polynucleotides [illegible] proteases,
PT   useful for treatment, diagnosis or prevention [illegible]
PT   [illegible] inflammatory [illegible]
PT   disorders.
XX
PS   [illegible]
XX
CC   The invention relates to human proteases (PRTS [illegible]) and [illegible] corresponding
CC   cDNA [illegible]
CC   the diagnosis, treatment and prevention of [illegible] with
CC   increased or decreased expression of PRTS. Examples of such disorders
CC   include cell proliferative disorders ([illegible]
CC   hepatitis [illegible] and cancers), autoimmune/inflammatory disorders
CC   (AIDS, Addison's disease, allergy, anaemia, asthma, atopic dermatitis,
CC   diabetes mellitus, glomerulonephritis, multiple sclerosis, osteoporosis,
CC   [illegible] disease, rheumatoid arthritis [illegible] and
CC   viral, bacterial, fungal, parasitic [illegible]
CC   infections); cardiovascular disorders (myocardial infarction, ischaemic
CC   heart disease and hypertension); neurological disorders (epilepsy,
CC   [illegible]
CC   [illegible] disorders, stroke, mental disorders [illegible] anxiety,
CC   and seasonal affective disorder and prion diseases); gastrointestinal
CC   disorders [illegible]
CC   [illegible]
CC   [illegible] insect bites and urticaria); reproductive disorders
CC   [illegible] and
CC   [illegible]
CC   [illegible]
CC   [illegible] cDNA [illegible] PRTS and [illegible]
CC   [illegible]
CC   [illegible] the present sequence [illegible] human protease [illegible] cDNA.
XX
SQ   Sequence [illegible]

Alignment Scores:
Pred. No.:        [illegible]    Length:        [illegible]
Score:            [illegible]    Matches:       [illegible]
Percent Similarity: [illegible]  Conservative:  [illegible]
Best Local Similarity: [illegible] Mismatches:  [illegible]
Query Match:      [illegible]    Indels:        [illegible]
DB:               [illegible]    Gaps:          [illegible]

US-10-070-464-1 (1-[illegible]) x AAI[illegible] (1-[illegible])

Qy   [illegible]
Db   [illegible]
```

```
XX
XX
CC   The present invention relates to [illegible]
CC   proteases referred to as dipeptidyl peptidase IV (DPPIV) related
CC   proteins (DRP). The dipeptidyl peptidase IV related proteins (DRP)
CC   and nucleic acids encoding them are useful for treating infections
CC   such as fungal, bacterial, protozoan and viral infections, particularly
CC   infections caused by human immunodeficiency virus (HIV-1 or HIV-2),
CC   pain, diabetes, precocious puberty, infertility, obesity, anorexia,
CC   bulimia, Parkinson's disease, acute heart failure, hypotension,
CC   hypertension, urinary retention, osteoporosis, angina pectoris,
CC   stroke, [illegible], asthma, allergies, [illegible]
CC   psychotic and neurological disorders ([illegible] dementia, [illegible]
CC   schizophrenia), and dyskinesias. [illegible] may also be used in [illegible]
CC   [illegible]
CC   [illegible] AeK[illegible] [illegible] DRP proteins.
XX
SQ   Sequence [illegible] 1424 [illegible]

Alignment Scores:
Pred. No.:             [illegible]      Length:        [illegible]
Score:                 [illegible]      Matches:       [illegible]
Percent Similarity:    94.22%           Conservative:  [illegible]
Best Local Similarity: 94.22%           Mismatches:    [illegible]
Query Match:           [illegible]      Indels:        [illegible]
DB:                    24               Gaps:          [illegible]

Qy          1 MetAlaAlaAlaMetGluThr[illegible]
Db        214 ATGGCA[illegible]

Qy         21 [illegible]
Db        274 GAGGAGAA[illegible]
```

[illegible]

RESULT [illegible]
AAK[illegible]
ID AB[illegible]
XX
AC AB[illegible]
XX
DT [illegible]
XX
DE cDNA [illegible] splice variant #[illegible]
XX
KW Human [illegible] dipeptidyl peptidase IV related [illegible]
KW DPP-IV; [illegible] human immunodeficiency virus; HIV-1; HIV-2; [illegible]
KW diabetes; [illegible] obesity; anorexia; Parkinson's disease; [illegible]
KW heart failure; [illegible] urinary retention; osteoporosis; [illegible]
KW ulcer; [illegible] psychotic disorder; neurological disorder;
KW dyskinesia; [illegible] disorder; inflammatory disorder;
KW metal [illegible]
XX
OS Homo sapiens.
XX
PN W[illegible]
XX
PD [illegible]
XX
PF [illegible]
XX
PR [illegible]
XX
PA ([illegible]) [illegible]
XX
PI [illegible]
XX
DR WPI; [illegible]
DR P-PSDB; AB[illegible]
XX
PT New [illegible] peptidase IV related [illegible]
PT the [illegible] treating [illegible] bacterial [illegible]
PT [illegible]

```
Qy    [illegible]

Db    [illegible]

RESULT [illegible]
ABK[illegible]
ID   ABK[illegible] standard; cDNA; [illegible] BP.
XX
AC   ABK[illegible];
XX
DT   [illegible] (first entry)
XX
DE   cDNA encoding human DPRP [illegible] splice variant [illegible].
XX
KW   Human; serine protease; dipeptidyl peptidase IV-related protein; [illegible]
KW   DPRP; infection; human immunodeficiency virus; HIV [illegible] HIV [illegible]
KW   diabetes; infertility; obesity; anorexia; Parkinson's disease; [illegible]
KW   heart failure; hypertension; urinary retention; osteoporosis; [illegible]
KW   ulcer; allergy; cancer; psychotic disorder; [illegible] disorder;
KW   dyskinesia; reproductive disorder; inflammatory disorder;
KW   metabolic disorder; [illegible].
XX
OS   Homo sapiens.
XX
PN   W[illegible] A.
XX
PD   [illegible]
XX
PF   [illegible] 2[illegible] 2[illegible] W[illegible].
XX
PR   [illegible] GB 2[illegible].
XX
PA   ([illegible]) [illegible] BV.
XX
PI   [illegible] S, Akinsanya K, [illegible], [illegible];
XX
DR   WPI; [illegible]-2-444[illegible]/47.
DR   [illegible]; ABK[illegible].
XX
PT   New dipeptidyl peptidase IV-related protein [illegible]
PT   the proteins, useful for treating e.g. fungal, bacterial, [illegible]
PT   viral infections, cancers, allergies, [illegible] disorders, [illegible]
PT
XX
PS   [illegible]
XX
CC   The present invention relates to the isolation of novel human serine
CC   proteases referred to as dipeptidyl peptidase IV (DPP-IV)-related
CC   proteins (DPRPs). The dipeptidyl peptidase IV-related proteins (DPRP)
CC   and nucleic acids encoding them are useful [illegible]
CC   such as fungal, bacterial, protozoan and viral infections [illegible]
CC   infections caused by human immunodeficiency virus (HIV-1 [illegible] HIV-2),
CC   pain, diabetes, precocious puberty, infertility, obesity, anorexia,
CC   bulimia, Parkinson's disease, acute heart failure, hypotension,
CC   hypertension, urinary retention, osteoporosis, [illegible]
CC   [illegible] asthma, allergies, [illegible]
CC   psychotic and neurological disorders [illegible]
CC   [illegible] and dyskinesias. These may also be used [illegible]
CC   therapeutic agents for the treatment of reproductive, inflammatory [illegible]
CC   metabolic disorders. ABK[illegible]22 ABK[illegible] human [illegible]
XX
SQ   Sequence 4[illegible]

Alignment Scores:
Pred. No.:            [illegible]     Length:        4[illegible]
Score:                4[illegible]    Matches:       [illegible]
Percent Similarity:   [illegible]     Conservative:  [illegible]
Best Local Similarity: [illegible]    Mismatches:    [illegible]
Query Match:          [illegible]     Indels:        [illegible]
DB:                   24              Gaps:          [illegible]

US-10-070-464-1 ([illegible]) x ABK[illegible] ([illegible] 4[illegible])
```

[illegible]

PA (MILL-) MILLENNIUM PHARM INC.

CC muscular dystrophy, epithelial [illegible] dermat[illegible] neurological [illegible]
CC epilepsy and Alzheimer's disease and [illegible]
CC [illegible]
CC [illegible]
CC [illegible]
XX
SQ [illegible]

Alignment Scores:
Pred. No.: [illegible] Length: [illegible]
Score: [illegible] Matches: [illegible]
Percent Similarity: [illegible] Conservative: [illegible]
Best Local Similarity: [illegible] Mismatches: [illegible]
Query Match: [illegible] Indels: [illegible]
DB: 24 Gaps: [illegible]

US-10-070-464-1 [illegible] x [illegible] [illegible]

Qy 1 MetAla[illegible]AlaMet[illegible]
Db [illegible]

Qy 21 GlnGlnAsnIleGlnSerGlnAspArgProLysLeu[illegible]ProPheLysVal[illegible]GluArgTyr [illegible]
Db [illegible]

Qy [illegible] [illegible]
Db [illegible]

Qy [illegible] [illegible]HisAsp[illegible]
Db [illegible]

Qy [illegible] AspArg[illegible]TyrLeuAlaMetSerGlyGlnAsnArg[illegible]
Db [illegible]

Qy [illegible] [illegible]
Db [illegible]

Qy [illegible] Asp[illegible]
Db [illegible]

Qy [illegible] [illegible]
Db [illegible]

Qy 161 [illegible]PheGlnAla[illegible]SerGly[illegible]
Db [illegible]

Qy [illegible] [illegible]
Db [illegible]

Qy [illegible] Asp[illegible]
Db [illegible]

Qy [illegible] [illegible]AsnIleVal[illegible]
Db [illegible]

Qy [illegible] [illegible]AsnMet[illegible]AspAla[illegible]
Db [illegible]

Qy [illegible] [illegible]

[illegible]

[illegible]

```
XX
KW   Human protease [illegible]
XX
FT                   /label= "Human protease [illegible]"
XX
PA   (INCY-) INCYTE [illegible]
XX
PI   [illegible] Kallick DA.
XX
PS   [illegible]
XX
CC   Twenty-one human proteases (referred to as HPRS [illegible]) [illegible]
CC   [illegible] (e.g. atherosclerosis [illegible]) and [illegible]
CC   [illegible] disorders.
XX
CC   The present [illegible] relates to twenty-one new human proteases
CC   [illegible]. The HPRS polynucleotides and
CC   [illegible]
CC   [illegible] gastrointestinal [illegible]
CC   [illegible]
CC   myocardial [illegible] / inflammatory [illegible] acquired
CC   immunodeficiency syndrome (AIDS), allergies and osteoarthritis [illegible]
```

```
OS   Homo sapiens.
XX
PN   W[illegible]4 A[illegible]
XX
PD   [illegible] APR 2002.
XX
PF   [illegible] 2001; [illegible]
XX
PR   [illegible]
XX
PA   (FERR-) FERRING BV.
XX
PI   [illegible] S, Akinsanya K, Riviere P[illegible]
XX
DR   WPI; [illegible]-444[illegible]/47.
DR   P-PSDB; ABG[illegible].
XX
PT   New dipeptidyl peptidase IV related proteins and nucleic acids encoding
PT   the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT   viral infections, cancers, allergies, [illegible] disorders [illegible]
PT
XX
PS   Claim [illegible]; [illegible] English.
XX
CC   The present invention relates to the [illegible] human serine
CC   proteases referred to as dipeptidyl peptidase IV (DPPIV) related
CC   proteins (DRPs). The dipeptidyl peptidase IV related proteins (DRPs)
CC   and nucleic acids encoding them are used [illegible] treating infections
CC   such as fungal, bacterial, protozoan and viral infections, particularly
CC   infections caused by human immunodeficiency virus (HIV-1 or HIV-2),
CC   pain, diabetes, prostatic puberty, infertility, obesity, anorexia,
CC   bulimia, Parkinson's disease, acute heart failure, hypotension,
CC   hypertension, urinary retention, osteoporosis, angina pectoris,
CC   strokes, ulcers, asthma, allergies, cancers, migraine, vertigo,
CC   psychotic and neurological disorders (e.g. anxiety, dementia [illegible]
CC   schizophrenia), and dyskinesias. These may also be used [illegible]
CC   therapeutic agents for the treatment of reproductive, inflammatory and
CC   metabolic disorders. ABK[illegible] ABK[illegible]4 encode human DRP proteins
XX
SQ   Sequence [illegible] BP; [illegible]

Alignment Scores:
Pred. No.:                            Length:
Score:                  4[illegible]             Matches:
Percent Similarity:                   Conservative:
Best Local Similarity:                Mismatches:
Query Match:                          Indels:
DB:                     24            Gaps:
```

[illegible]

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	[illegible]	[illegible]	[illegible]	21	ABK8[illegible]	cDNA [illegible]
2	[illegible]	[illegible]	[illegible]	24	ABK1[illegible]	Human [illegible]
3	[illegible]	[illegible]	[illegible]	22	AA[illegible]	Nucleotide [illegible]
4	[illegible]	[illegible]	[illegible]	21	AA[illegible]	Human [illegible]
5	[illegible]	[illegible]	[illegible]	21	AA[illegible]	cDNA [illegible]
6	[illegible]	[illegible]	[illegible]	24	AA[illegible]	Coding [illegible]
7	[illegible]	[illegible]	[illegible]	21	ABK8[illegible]	cDNA [illegible]
8	[illegible]	[illegible]	[illegible]	21	ABK8[illegible]	cDNA [illegible]
9	[illegible]	[illegible]	[illegible]	21	ABK8[illegible]	cDNA [illegible]
10	[illegible]	[illegible]	[illegible]	24	ABN[illegible]	Novel human [illegible]
11	[illegible]	[illegible]	[illegible]	21	ABK8[illegible]	cDNA [illegible]
12	[illegible]	[illegible]	[illegible]	21	AA[illegible]	Human [illegible]
13	[illegible]	[illegible]	[illegible]	24	ABN[illegible]	Novel human [illegible]
14	[illegible]	[illegible]	[illegible]	24	ABK8[illegible]	cDNA [illegible]
15	[illegible]	[illegible]	[illegible]	[illegible]	AAU[illegible]	Human cDNA [illegible]
16	[illegible]	[illegible]	[illegible]	21	ABK8[illegible]	cDNA [illegible]
17	[illegible]	[illegible]	[illegible]	21	ABK8[illegible]	cDNA [illegible]
18	[illegible]	[illegible]	[illegible]	21	ABK8[illegible]	cDNA [illegible]
19	[illegible]	[illegible]	[illegible]	21	AAD[illegible]	Human dipeptidyl [illegible]
20	[illegible]	[illegible]	[illegible]	21	AAD[illegible]	Human dipeptidyl [illegible]
21	[illegible]	[illegible]	[illegible]	21	AAD[illegible]	Alternative [illegible]
22	[illegible]	[illegible]	[illegible]	21	ABK8[illegible]	cDNA encoding [illegible]
23	[illegible]	[illegible]	[illegible]	24	ABK8[illegible]	cDNA encoding [illegible]
24	[illegible]	[illegible]	[illegible]	[illegible]	AAD[illegible]	Murine dipeptidyl [illegible]
25	[illegible]	[illegible]	[illegible]	21	ABK8[illegible]	cDNA encoding [illegible]
26	[illegible]	[illegible]	[illegible]	24	ABK8[illegible]	cDNA encoding [illegible]
27	[illegible]	[illegible]	[illegible]	22	AA[illegible]	Human [illegible]
28	[illegible]	[illegible]	[illegible]	24	ABK8[illegible]	cDNA encoding [illegible]
29	[illegible]	[illegible]	[illegible]	24	ABK8[illegible]	cDNA encoding [illegible]
30	[illegible]	[illegible]	[illegible]	22	AA[illegible]	Human [illegible]
31	[illegible]	[illegible]	[illegible]	22	AA[illegible]	Nucleotide sequence [illegible]
32	[illegible]	[illegible]	[illegible]	22	AA[illegible]	Human [illegible]
33	[illegible]	[illegible]	[illegible]	[illegible]	AA[illegible]	Human [illegible]
34	[illegible]	[illegible]	[illegible]	24	ABK6[illegible]	DNA encoding human [illegible]
35	[illegible]	[illegible]	[illegible]	21	ABK6[illegible]	DNA encoding human [illegible]
36	[illegible]	[illegible]	[illegible]	22	AA[illegible]	Nucleotide sequence [illegible]
37	[illegible]	[illegible]	[illegible]	22	AA[illegible]	Nucleotide sequence [illegible]
38	[illegible]	[illegible]	[illegible]	[illegible]	AA[illegible]	Human [illegible]
39	[illegible]	[illegible]	[illegible]	[illegible]	ABL[illegible]	Drosophila [illegible]
40	[illegible]	[illegible]	[illegible]	[illegible]	ABL[illegible]	Drosophila [illegible]
41	[illegible]	[illegible]	[illegible]	[illegible]	AAA[illegible]	Human [illegible]
42	[illegible]	[illegible]	[illegible]	[illegible]	ABK8[illegible]	cDNA [illegible]
43	[illegible]	[illegible]	[illegible]	[illegible]	AA[illegible]	Human [illegible]
44	[illegible]	[illegible]	[illegible]	[illegible]	ABL[illegible]	Drosophila [illegible]
45	[illegible]	[illegible]	[illegible]	[illegible]	ABL[illegible]	Drosophila [illegible]

get

ALIGNMENTS

RESULT 1
ABK8[illegible]
ID ABK8[illegible]
XX
AC ABK8[illegible]
XX
DT [illegible]
XX
DE cDNA encoding human [illegible] related serine protease [illegible]
XX
KW Human; serine protease; dipeptidyl peptidase IV related protease; [illegible]
KW [illegible] human immunodeficiency virus; HIV [illegible]
KW diabetes; [illegible] obesity; anorexia; Parkinson's disease; [illegible]
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; [illegible] psychiatric disorder; neurological disorder;
KW dyskinesia; [illegible] disorder; inflammatory disorder;
KW metabolic [illegible]
XX

```
[illegible] AA AA AA [illegible]

RESULTS [illegible]
A [illegible]
LOCUS        A [illegible]  NA  [illegible]  PRI [illegible]
DEFINITION   Homo sapiens chromosome 17, [illegible] complete [illegible]
ACCESSION    A [illegible]
VERSION      A [illegible]
KEYWORDS     HTG.
SOURCE       Homo sapiens.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; [illegible] Vertebrata; [illegible]
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to [illegible])
  AUTHORS    Birren,B., [illegible], Nusbaum,C. and Lander,E.
  TITLE      Homo sapiens chromosome 17, clone [illegible]
  JOURNAL    Unpublished
REFERENCE    2  (bases 1 to [illegible])
  AUTHORS    Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Abb[illegible]
             Baker,[illegible], Baldwin,[illegible], Barna,N., [illegible], Bennett,[illegible]
             Castle,A., Cerny,L., Colangelo,M., [illegible]
             [illegible], DeArellano,K., [illegible], Dewar,K.,
             [illegible]
             [illegible]
             [illegible]
             Karatas,A., Lehoczky,J., Lieu,[illegible], MacDonald,P.,
             Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,[illegible]
             Meldrim,J., Meyer,M., Morris,W., [illegible]
             Naylor,J., Nicol,M., O'Connor,[illegible], Farlin,[illegible]
             Peterson,K., Rollins,V., Riley,[illegible], Roberts,[illegible], Roy,A., S[illegible]
             Stange-Thomann,N., Stojanovic,N., [illegible], Subramanian,A.,
             Tesfaye,S., Vermuelen Miller,J., Vassiliev,H., Vo,A., Wamer,A.,
             Wheeler,[illegible], Wu,X., Wyman,D., Y[illegible], [illegible]
  TITLE      Direct Submission
  JOURNAL    Submitted (22-MAY-19[illegible]) Whitehead Institute/MIT Center for Genome
             Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE    3  (bases 1 to 94205)
  AUTHORS    Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Abb[illegible]
             Baker,[illegible], Baldwin,[illegible], Barna,N., [illegible], Benn[illegible]
             Castle,A., Cerny,[illegible], Colangelo,M., Collins,[illegible], [illegible]more,A.
             [illegible], DeArellano,K., [illegible], Dewar,K.,
             [illegible]
             [illegible]
             [illegible]
             Karatas,A., Lehoczky,J., Lieu,[illegible], MacDonald,P.,
             Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,[illegible]
             Meldrim,J., Meyer,M., Morris,W., [illegible]
             Naylor,J., Nicol,M., O'Connor,[illegible], [illegible], Farlin,[illegible]
             Peterson,K., Rollins,V., Riley,[illegible], Roberts,[illegible], Roy,A., [illegible]
             Stange-Thomann,N., Stojanovic,N., [illegible], Subramanian,A.,
             Tesfaye,S., Vermuelen Miller,J., Vassiliev,H., Vo,A., Wamer,A.,
             Wheeler,[illegible], Wu,X., Wyman,D., [illegible]
  TITLE      Direct Submission
  JOURNAL    Submitted (1[illegible]) Whitehead [illegible]
             Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT      On [illegible] this sequence version replaced gi:[illegible]
             All repeats were identified using RepeatMasker: [illegible]
             Green, P. (1996-1997)
             http://ftp.genome.washington.edu/RM/RepeatMasker.html

             Only [illegible] base pairs from the [illegible]
             [illegible] The remaining [illegible]
             (WIBR project [illegible]) or accession number A[illegible] (W[illegible]
             [illegible]).
FEATURES             Location/Qualifiers
     source          1..[illegible]
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="[illegible]"
                     /clone_lib="[illegible] Library"
```

```
repeat_region   [illegible]
repeat_region   [illegible]
repeat_region   [illegible]
[illegible]
```

```
AUTHORS   [illegible]
TITLE     Direct Submission
JOURNAL   Submitted (24-[illegible]) [illegible]
          Genomics Laboratory; [illegible] Yana, Kisarazu, Chiba [illegible]
          [illegible]
COMMENT   NEDO human cDNA sequencing project supported by Ministry of
          Economy, Trade and Industry of Japan; cDNA full insert sequence:
          Research Association for Biotechnology (RAB); cDNA library
          construction: Helix Research Institute (HRI) (supported by [illegible]
          Key Technology Center [illegible]); [illegible] sequencing:
          HRI, and Biotechnology Center, National Institute of [illegible]
          [illegible]
          HRI.
FEATURES             Location/Qualifiers
     source          [illegible]
                     /organism="Homo sapiens"
                     /db_xref="taxon:[illegible]"
                     /clone="HES[illegible]"
                     /clone_lib="[illegible]"
                     /clone_lib="HES[illegible]"
                     /note="[illegible]"
BASE COUNT    [illegible]
ORIGIN

Alignment Scores:
Pred. No.:            [illegible]     Length:       [illegible]
Score:                [illegible]     Matches:      [illegible]
Percent Similarity:   [illegible]     Conservative: [illegible]
Best Local Similarity: [illegible]    Mismatches:   [illegible]
Query Match:          [illegible]     Indels:       [illegible]
DB:                   [illegible]     Gaps:         [illegible]

SW [illegible] (1-7) x AK0[illegible] ([illegible])

Qy       [illegible]V[illegible]
               [illegible]
Db       [illegible]AA[illegible]AA[illegible]

RESULT [illegible]
AX[illegible]
LOCUS       AX[illegible]     [illegible] bp    DNA     linear   PAT [illegible]
DEFINITION  Sequence [illegible] from Patent WO[illegible].
ACCESSION   AX[illegible]
VERSION     AX[illegible]  GI:[illegible]
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yue,H., [illegible],V.[illegible], [illegible],A.[illegible], [illegible],
            [illegible],M., [illegible],A.M., [illegible], Nguyen,D.[illegible],
            [illegible],A., Khan,F.A., Walia,N.K., [illegible], [illegible],
            [illegible], Walsh,R.[illegible], [illegible], [illegible] and [illegible]
  JOURNAL   Patent: WO [illegible]
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..[illegible]
                     /organism="Homo sapiens"
                     /db_xref="taxon:[illegible]"
                     /note="Incyte ID No: [illegible]"
BASE COUNT   [illegible]
ORIGIN

Alignment Scores:
Pred. No.:            [illegible]     Length:       [illegible]
Score:                [illegible]     Matches:      [illegible]
Percent Similarity:   [illegible]     Conservative: [illegible]
Best Local Similarity: [illegible]    Mismatches:   [illegible]
Query Match:          [illegible]     Indels:       [illegible]
DB:                   [illegible]     Gaps:         [illegible]

SW [illegible] x AX[illegible] ([illegible])
```

```
Qy       [illegible]
Db       [illegible]

RESULT [illegible]
AK[illegible]
LOCUS       AK[illegible]
DEFINITION  [illegible]
ACCESSION   AK[illegible]
VERSION     AK[illegible]
KEYWORDS    [illegible]
SOURCE      [illegible]
  ORGANISM  [illegible]
REFERENCE   [illegible]
  AUTHORS   [illegible]
  TITLE     [illegible]
  JOURNAL   [illegible]
  MEDLINE   [illegible]
REFERENCE   [illegible]
  AUTHORS   [illegible]
  TITLE     Direct Submission
  JOURNAL   [illegible]
FEATURES             Location/Qualifiers
     source          [illegible]
     CDS             [illegible]
                     /translation="[illegible]"
BASE COUNT   [illegible]
ORIGIN

Alignment Scores:
Pred. No.:            [illegible]
Score:                [illegible]
Percent Similarity:   [illegible]
Best Local Similarity: [illegible]
Query Match:          [illegible]
```

[illegible]

TITLE A new testis development gene NYD-[illegible] from human testis

[illegible]

JOURNAL Submitted ([illegible]) [illegible] Medical University, Han Zhong Road 140, Nanjing, Jiangsu 210029, China

[illegible]

FEATURES [illegible]

/translation="[illegible]"

[illegible]

ORIGIN

Alignment [illegible]

[illegible]

FEATURES [illegible]

/organism="Homo sapiens"

/db_xref="taxon:9606"

/translation="MAAAMP[illegible]"

[illegible]

ORIGIN

Alignment [illegible]

[illegible]

DEFINITION Homo sapiens cDNA FLJ[illegible] fis, clone TESTI20[illegible]

ACCESSION AK0[illegible]

VERSION AK0[illegible]

KEYWORDS [illegible] (full insert sequence).

SOURCE Homo sapiens testis cDNA to mRNA, clone TESTI20[illegible]

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS [illegible] Matsumoto,K., Hirano,M., Sano,S., Nomura,[illegible] Matsunawa,H., [illegible] Watanabe,M., Sugiyama,[illegible] Yamamoto,[illegible] Nishikawa,[illegible] Kimura,K., [illegible] Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., [illegible] Kawakami,B., [illegible] and [illegible]

TITLE [illegible] sequencing project

JOURNAL Unpublished

REFERENCE 2 [illegible]

ACCESSION AX4...
VERSION AX4...
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Novel nucleic acids and polypeptides
FEATURES Location/Qualifiers
/organism="Homo sapiens"
/note="unnamed protein product"
/protein_id=
/translation="MAAAMEE...

/organism="Homo sapiens"
DEFINITION Homo sapiens cDNA FLJ...
ACCESSION AK...
VERSION AK...
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Kawakami...

```
SOURCE      Homo sapiens
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
            Mammalia; Eutheria; Primates; [illegible]
REFERENCE   1  (bases 1 to [illegible])
  AUTHORS   Abbott,C.A., Yu,D.M., Woollatt,E., Sutherland,G.R., McCaughan,G.W.
            and Gorrell,M.D.
  TITLE     [illegible] expression and chromosomal [illegible]
            dipeptidyl peptidase (DPP) IV [illegible]
  MEDLINE   [illegible]
REFERENCE   2  (bases 1 to [illegible])
  AUTHORS   Abbott,C.A., Yu,D., McCaughan,G.W. and Gorrell,M.D.
  TITLE     Direct Submission
  JOURNAL   Submitted ([illegible]-JAN-[illegible]) A.W. Morrow [illegible]
            Centre, Centenary Institute of [illegible] and [illegible]
            [illegible] NSW [illegible] Australia
FEATURES             Location/Qualifiers
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /gene="[illegible]"
                     /gene="[illegible]"
                     /note="[illegible]
                     variants of this mRNA are found in GenBank Accession
                     Numbers [illegible]"
                     /codon_start=1
                     /product="dipeptidyl peptidase [illegible]"
                     /translation="[illegible]
                     [illegible]
                     QSSGGRVAIAGALVLWIFVI[illegible]"
BASE COUNT  [illegible]
ORIGIN

Alignment Scores:
[illegible]
```

SUMMARY

look

```
FEATURES             [illegible]
                     [illegible]
                     /organism="[illegible]"
                     /[illegible]="[illegible]"
                     /clone="[illegible]"
                     /clone_lib="[illegible]"
                     /tissue_type="[illegible]"
                     /lab_host="[illegible] (phage-resistant)"
                     /note="[illegible] Vector: [illegible]
                     model: cDNA made by oligo-dT primer. [illegible]
                     [illegible] EcoRI/XhoI [illegible]
                     adaptor: [illegible]. Library constructed by [illegible]
                     in the laboratory of [illegible] (University of
                     California, Berkeley) [illegible] cDNA synthesis kit
                     (Stratagene) and [illegible]
                     Note: this is a NIH MGC Library."
BASE COUNT     [illegible]
ORIGIN

Alignment Scores:
[illegible]               [illegible]          Length:        [illegible]
Score:                    [illegible]          Matches:       [illegible]
Percent Similarity:       [illegible]          Conservative:  [illegible]
Best Local Similarity:    [illegible]          Mismatches:    [illegible]
Query Match:              [illegible]          Indels:        [illegible]
DB:                       14                   Gaps:          [illegible]

[illegible]

Qy     [illegible]
           [illegible]
Db     [illegible] AA AA AA [illegible]


Search completed: February [illegible], 2003, [illegible]
[illegible]
```

```
LOCUS       BG7[illegible]0182                  [illegible]    mRNA    linear   EST [illegible]
DEFINITION  [illegible] NIH_MGC_[illegible] Homo sapiens cDNA [illegible] IMAGE:[illegible],
            mRNA sequence.
ACCESSION   BG7[illegible]82
VERSION     [illegible]
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
            Mammalia; Eutheria; Primates; [illegible]
REFERENCE   1  (bases 1 to 85[illegible])
  AUTHORS   NIH MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: [illegible]
            cDNA Library Preparation: [illegible] (NHGRI) [illegible]
            [illegible] and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: [illegible]
            [illegible]
            [illegible] I.M.A.G.E. Consortium (LLNL) at:
            http://image.llnl.gov
            [illegible]
            High quality sequence [illegible]
FEATURES             Location/Qualifiers
     source          1..[illegible]
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:[illegible]"
                     /clone_lib="NIH_MGC_[illegible]"
                     /tissue_type="[illegible]"
                     /note="Organ: testis; Vector: [illegible]
                     [illegible]
```

[illegible]

```
DEFINITION  [illegible] XS [illegible] mRNA [illegible]
            mRNA sequence.
ACCESSION   [illegible]
VERSION     [illegible]
KEYWORDS    EST.
SOURCE      [illegible]
  ORGANISM  Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
            Amphibia; Batrachia; Anura; Mesobatrachia; [illegible]; Pipidae;
            Xenopodinae; Silurana.
REFERENCE   1  (bases 1 to [illegible])
  AUTHORS   Taylor,R., Ashurst,J.L., [illegible], Zorn,A.M. and [illegible]
  TITLE     Sanger Xenopus tropicalis EST project [illegible]
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Taylor R
            Sanger Centre
            [illegible] Cambridgeshire, [illegible]
            Email: [illegible]
            Sanger Xenopus tropicalis EST project [illegible]
            [illegible] SEQUENCE ID: [illegible]
            Sequencing primer: [illegible]
            This sequence is from a Xenopus [illegible] (XTE) [illegible]
            constructed by Aaron M. Zorn.
FEATURES             Location/Qualifiers
     source          [illegible]
                     /organism="Silurana tropicalis"
                     /db_xref="taxon:[illegible]"
                     /clone="[illegible]"
                     /clone_lib="XTE neurula"
                     /dev_stage="neurula"
                     /lab_host="Escherichia coli [illegible]"
                     /note="Vector: pCS[illegible]; Site_1: [illegible]; Site_2: [illegible]
                     was [illegible] primed [illegible] RNA [illegible]
                     [illegible] Not I [illegible] cDNA was [illegible] cloned into pCS[illegible] with
                     [illegible] at the 5' end and Not I at the 3' end."
BASE COUNT  [illegible]
ORIGIN
```

[illegible]

[illegible]

SUMMARIES

ALIGNMENTS

[illegible]

The present invention relates to [illegible] proteases referred to as dipeptidyl peptidase IV (DPPIV) related protease (DRP). The dipeptidyl peptidase IV related protease (DRP) and nucleic acids encoding them are useful for treating infections such as fungal, bacterial, protozoan and viral infections, particularly infections caused by human immunodeficiency virus (HIV 1 or HIV 2), pain, diabetes, precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, angina pectoris, stroke, [illegible] asthma, allergies, cancers, migraines, vomiting, psychotic and neurological disorders (e.g. [illegible]), and dyskinesias. These may also be used in [illegible] for the treatment of reproductive, inflammatory [illegible] disorders. [illegible] encode human [illegible] proteins.

[illegible]

XX

XX

The present invention relates to [illegible] proteases referred to as dipeptidyl peptidase IV (DPPIV) related protease (DRP). The dipeptidyl peptidase IV related protease (DRP) and nucleic acids encoding them are useful for treating infections such as fungal, bacterial, protozoan and viral infections, particularly infections caused by human immunodeficiency virus (HIV 1 or HIV 2), pain, diabetes, precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, angina pectoris, stroke, [illegible] asthma, allergies, cancers, migraines, vomiting, psychotic and neurological disorders (e.g. anxiety, dementia, schizophrenia), and dyskinesias. These may also be used in [illegible] for the treatment of reproductive, inflammatory [illegible] metabolic disorders. [illegible] encode human DRP proteins.

XX

Alignment [illegible]

[illegible]

```
XX
DR   W[illegible]/4[illegible].
DR   [illegible] AAP[illegible]
XX
PT   New dipeptidyl peptidase (DPP) peptides, useful [illegible]
PT   inhibitors of DPP catalytic activity, which [illegible]
PT   [illegible] neoplasia, type II diabetes, cirrhosis, autoimmunity [illegible]
PT   rejection and HIV infection.
XX
PS   Example; [illegible]
XX
CC   The present invention relates to dipeptidyl peptidase (DPP) prote[illegible]
CC   polynucleotides encoding such proteins. The [illegible] peptides are useful [illegible]
CC   screening inhibitors of DPP catalytic activity. The inhibitors are [illegible]
CC   for treating neoplasia, type II diabetes, cirrhosis, autoimmunity [illegible]
CC   rejection and HIV (human immune deficiency virus) infection. The prese[illegible]
CC   sequence is human DPP[illegible] cDNA.
XX
SQ   Sequence [illegible] BP; [illegible] A; [illegible] C; [illegible] G; [illegible] T; [illegible] other;

Alignment Scores:
Pred. No.:              [illegible]      Length:      [illegible]
Score:                  [illegible]      Matches:     [illegible]
Percent Similarity:     100.0%           Conservative: [illegible]
Best Local Similarity:  [illegible]      Mismatches:  [illegible]
Query Match:            [illegible]      Indels:      [illegible]
DB:                     24               Gaps:        [illegible]

SW [illegible] (1-[illegible]) x AA[illegible] (1-[illegible])

[illegible]
[illegible]
[illegible] AAAAAA[illegible]

RESULT [illegible]
AA[illegible]
ID   AA[illegible] standard; cDNA; [illegible] BP.
XX
AC   AA[illegible];
XX
DT   [illegible] (first entry)
XX
DE   cDNA encoding [illegible] human prolyl [illegible] peptidase.
XX
KW   [illegible] prolyl oligopeptidase; [illegible]; endopeptidase;
KW   [illegible]
KW   [illegible] disorder; vascular disorder; [illegible] disorder; [illegible]
KW   antidiabetic; antiarthritic; antiasthmatic; anti-inflammatory;
KW   diabetes mellitus; arthritis; [illegible] asthma;
KW   Crohn's disease; neuronal disorder; demyelination disease; [illegible]
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             [illegible]..28[illegible]
FT                   /*tag= [illegible]
FT                   /product= "[illegible] oligopeptidase"
FT                   /note= "This region is specifically claimed in
FT                   claim 2"
XX
PN   W[illegible]
XX
PD   [illegible]
XX
PF   [illegible]
XX
PR   [illegible]
XX
PA   (M[illegible]) M[illegible] PHARM INC.
XX
PI   [illegible] Williamson M;
XX
DR   WPI; [illegible]/4[illegible].
```

```
XX
PT   New [illegible]
[illegible]
XX
CC   [illegible]
XX
SQ   [illegible]

Alignment Scores:
[illegible]

RESULT [illegible]
AA[illegible]
ID   AA[illegible]
XX
AC   AA[illegible]
XX
DT   [illegible]
XX
DE   [illegible]
XX
KW   [illegible]
XX
OS   [illegible]
XX
PN   W[illegible]
XX
PD   [illegible]
XX
```

[illegible]

XX

[illegible]

Human dipeptidyl peptidase 8 (DPP8) cDNA.

XX

[illegible]

[illegible]

XX

KW Human; antiinflammatory; vulnerary; [illegible]

KW [illegible] antiparkinsonian; protein therapy; [illegible]

KW expressed sequence [illegible]

XX

OS Homo sapiens.

XX

PA (HYSE-) HYSEQ INC.

XX

[illegible]

XX

XX

XX

[illegible]

[illegible]

XX
DE human protease PRTS-2 [illegible]
XX
KW [illegible]
XX
FH Key Location/Qualifiers
FT [illegible]
FT /product= "[illegible] protein"
XX
XX
PA (INCY) INCYTE GENOMICS INC.
XX
PI [illegible]
XX
DR WPI; [illegible]
DR P-PSDB; [illegible]
XX
PT [illegible]
XX
CC [illegible]

[illegible]

[illegible]

RESULT [illegible]

[illegible]

XX

[illegible]

KW Human [illegible]

XX

[illegible]

XX

PI New [illegible] diagnosis [illegible] gene mapping, identification [illegible] response [illegible] biodiversity.

XX

CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for [illegible] and gene mapping [illegible] in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques [illegible] (II) [illegible] to treat disease states [illegible] (II). (II) [illegible] for generating antibodies against [illegible] quantitating [illegible] in tissues, as molecular weight markers [illegible] and [illegible] patterns are useful [illegible] (I) and (II) are useful [illegible] disorders [illegible] aberrant protein expression or biological [illegible] sequences have applications [illegible] gene mapping, identification [illegible] response [illegible] disorders or other traits [illegible] and [illegible] products dependent on DNA and [illegible] AAF[illegible] represent [illegible] the invention.

Note: The sequence [illegible] this patent did not appear in the [illegible] specification [illegible] format [illegible] sequences.

XX

[illegible]

Alignment Scores: [illegible]

```
XX
PR   [illegible]
PR   27-AUG-[illegible]
PR   [illegible]
PR   [illegible] MAY [illegible]
PR   [illegible]
XX
PA   (HELI-) HELIX RES INST.
XX
PI   [illegible] Nishikawa [illegible]
PI   [illegible] Wakamatsu A, [illegible]
XX
DR   WPI; [illegible]
XX
PT   [illegible]
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs
XX
PS   Claim 1; SEQ ID NO [illegible]
XX
CC   The present invention describes primer sets for synthesising
CC   full-length cDNAs defined in the specification, where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide [illegible]
CC   [illegible] the complementary strand of a polynucleotide which comprises
CC   [illegible] the nucleotide sequences defined in the specification, where [illegible]
CC   oligonucleotide comprises at least 15 nucleotides; or (b) [illegible]
CC   [illegible] an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary [illegible]
CC   polynucleotide which comprises a [illegible] sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the [illegible]
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers [illegible]
CC   cDNAs easily without any specialised [illegible] AAH[illegible] AAH[illegible]
CC   AAH[illegible] AAH[illegible] represent human cDNA sequences; AAH[illegible]
CC   AAH[illegible] represent human amino acid sequences and AAH[illegible]
CC   represent [illegible]
CC   [illegible] the present invention.
XX
SQ   Sequence [illegible]

Alignment Scores:
Pred. No.:              [illegible]        Length:          [illegible]
Score:                  [illegible]        Matches:         [illegible]
Percent Similarity:     [illegible]        Conservative:    [illegible]
Best Local Similarity:  [illegible]        Mismatches:      [illegible]
Query Match:            [illegible]        Indels:          [illegible]
DB:                     [illegible]        Gaps:            [illegible]

[illegible] AAH[illegible]

Qy      [illegible]

Db      [illegible] AAGAAGAA[illegible]

RESULT [illegible]
AAH[illegible]
ID   AAH[illegible] standard; cDNA; [illegible]
XX
AC   AAH[illegible];
XX
DT   [illegible] (first entry)
XX
DE   Nucleotide sequence encoding human [illegible]
XX
KW   Human; dipeptidyl aminopeptidase [illegible]
KW   dipeptidyl peptidase [illegible]
```

```
KW   [illegible]
KW   [illegible]
KW   [illegible]
XX
OS   [illegible]
XX
FH   [illegible]
FT   [illegible]
FT   [illegible]
FT   [illegible]
FT   [illegible]
XX
PN   WO[illegible]
XX
PD   [illegible]
XX
PF   [illegible]
XX
PR   [illegible]
PR   [illegible]
XX
PA   [illegible]
XX
PI   [illegible]
XX
DR   WPI; [illegible]
DR   [illegible]
XX
PT   [illegible]
PT   [illegible]
PT   [illegible]
XX
PS   [illegible]
XX
CC   [illegible]
CC   dipeptidyl aminopeptidase [illegible]
CC   [illegible] cDNA [illegible]
CC   prolyl oligopeptidase [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
XX
SQ   [illegible]

Alignment Scores:
Pred. No.:              [illegible]
Score:                  [illegible]
Percent Similarity:     [illegible]
Best Local Similarity:  [illegible]
Query Match:            [illegible]
DB:                     [illegible]

[illegible] AAH[illegible]

Qy      [illegible]

Db      [illegible] AAGAA[illegible]

RESULT [illegible]
AAH[illegible]
ID   AAH[illegible]
XX
AC   AAH[illegible]
XX
```

SUMMARIES

Cooh

ALIGNMENTS

PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI [illegible]
XX
PS Claim [illegible]
XX
CC The invention relates to an isolated [illegible]
CC human secreted protein (II). (I) and (II) [illegible]
CC [illegible]
CC symptoms associated with the disorders [illegible]
CC [illegible]
CC Disorders which are diagnosed or treated [illegible] autoimmune diseases
CC [illegible]
XX
SQ Sequence [illegible] AA;

[illegible]

[illegible]

XX

[illegible] ... enzyme linked immunosorbent assays [illegible]. Disorders which [illegible] diagnosed or treated include autoimmune diseases [illegible] ... disorders [illegible] of the breast [illegible], cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, and [illegible] system disorders [illegible] Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal [illegible] ... can also be used to aid wound [illegible] ...

XX

[illegible]

```
XX
FA   [illegible] UNIV [illegible]
XX
[illegible]
DR   N [illegible] AA[illegible]
XX
[illegible]
XX
CC   [illegible] in AA[illegible] represent fragments of human
CC   dipeptidyl aminopeptidase [illegible]. [illegible] substrate specificity [illegible]
CC   H-Ala-Pro-pNA, H-Gly-Pro-pNA and H-Arg-Pro-pNA. Therefore, it is a
CC   [illegible] peptidase and a dipeptidyl [illegible]
CC   of hydrolysing the peptide bond [illegible]
CC   [illegible] with human [illegible] IV. [illegible]
CC   [illegible] substrate, and for [illegible]
CC   in [illegible] measuring the level of [illegible] gene expression in a [illegible]
CC   [illegible] expression [illegible]
CC   RNA in the cell. It is also useful for identifying [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible] dependent [illegible]
CC   [illegible]
CC   [illegible] suppressants.
XX
SQ   Sequence [illegible] AA;
[illegible]
XX
KW   [illegible]
KW   [illegible] disorders [illegible]
KW   [illegible] HIV [illegible]
XX
[illegible]
```

[illegible]

Look

[illegible]

Best Local Similarity [illegible] Pred. No. [illegible]
Matches [illegible] Conservative [illegible] Mismatches [illegible]

Qy [illegible]

Db [illegible] FHSVF [illegible]

RESULT [illegible]
[illegible]
hypothetical protein [illegible] Caenorhabditis elegans
OS Species: Caenorhabditis elegans
[illegible]
[illegible] Bradshaw [illegible]
submitted to the EMBL Data Library, [illegible]
A;Description: The sequence of C. elegans [illegible]
A;Reference number: [illegible]
A;Accession: [illegible]
A;Status: preliminary; translated from GB/EMBL/[illegible]
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: EMBL:[illegible]; PID:[illegible]
A;Experimental source: strain [illegible]
[illegible]
A;Gene: [illegible]
A;Map position: [illegible]
A;Introns: [illegible]

Query Match [illegible] Score [illegible]
Best Local Similarity [illegible] Pred. No. [illegible]
Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible]

Qy [illegible]

Db [illegible] FHSVF [illegible]

RESULT [illegible]
[illegible]
branched-chain amino acid aminotransferase [illegible]
OS Species: Agrobacterium tumefaciens
[illegible]
[illegible]
[illegible] Williams [illegible]
Science [illegible]
[illegible]
A;Reference number: [illegible]
A;Accession: [illegible]
A;Status: preliminary
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: GB:[illegible]; PID:[illegible]
[illegible]
A;Gene: [illegible]
A;Map position: [illegible]
[illegible] branched-chain amino acid aminotransferase

Query Match [illegible]
Best Local Similarity [illegible] Pred. No. [illegible]
Matches [illegible] Conservative [illegible] Mismatches [illegible]

Qy [illegible]

Db [illegible]

RESULT [illegible]
[illegible]
[illegible]
N;Alternate names: [illegible]
OS Species: Mus musculus (house mouse)

[illegible]

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible]

Qy [illegible]

Db [illegible]

RESULT [illegible]
[illegible]
hypothetical protein [illegible]
OS Species: [illegible]
[illegible]

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible]

Qy [illegible]

Db [illegible]

RESULT [illegible]
[illegible]

```
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible]
RT   "[illegible]
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science [illegible].
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   [illegible]
RT   "[illegible]"
RL   Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   [illegible]
RT   "[illegible]"
RL   Submitted (JUN [illegible]) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; [illegible]; AA[illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]; WD40 family 2.
DR   [illegible]
DR   SMART; SM[illegible]
DR   SMART; SM[illegible]; [illegible] like; 1.
DR   SMART; SM[illegible]
DR   PROSITE; PS[illegible]; NKN WN [illegible]
DR   PROSITE; PS[illegible]
KW   [illegible]
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible]

[illegible]

[illegible]

[illegible]
```

```
  Query Match          [illegible];  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy          1 ERHSIR 6
              [illegible]
Db        [illegible] ERHSIR [illegible]


RESULT 10
Q9HBM5
ID   Q9HBM5      PRELIMINARY;      PRT;   882 AA.
AC   Q9HBM5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Dipeptidyl peptidase 8.
GN   DPP8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20467134; PubMed=11012666;
RA   Abbott C.A., Yu D.M.T., Woollatt E., Sutherland G.R., McCaughan G.W.,
RA   Gorrell M.D.;
RT   "Cloning, expression and chromosomal localization of a novel human
RT   dipeptidyl peptidase (DPP) IV homolog, DPP8.";
RL   Eur. J. Biochem. 267:6140-6150(2000).
DR   EMBL; AF221634; AAG29766.1; -.
DR   MEROPS; S09.018; -.
DR   InterPro; IPR002469; DPPIV_N_term.
DR   InterPro; IPR001375; Peptidase_S9.
DR   InterPro; IPR002379; Ser_est[illegible]_site.
DR   Pfam; PF00930; DPPIV_N_term; 1.
DR   Pfam; PF00326; Peptidase_S9; 1.
SQ   SEQUENCE   882 AA;  101421 MW;  AB801[illegible] CRC64;

  Query Match          100.0%;  Score 31;  DB 1;  Length 882;
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches   6;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 ERHSIR 6
              ||||||
Db        647 ERHSIR 652


RESULT 11
Q9D4J9
ID   Q9D4J9      PRELIMINARY;      PRT;   882 AA.
AC   Q9D4J9;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   4933432P09RIK protein.
GN   4933432P09RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=TESTIS;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
```

```
RA   [illegible]
RA   Sakai K., [illegible]
RA   Rocke [illegible]
RA   Brusnstein M.J. [illegible]
RA   Dastgheib S., [illegible]
RA   [illegible]
RA   Nadamoe [illegible]
RA   Sasaki D., [illegible]
RA   Suzuki H., [illegible]
RA   Wynshaw-Boris A., [illegible]
RA   Hayashizaki Y.;
RT   "Functional annotation [illegible]";
RL   Nature [illegible]
DR   EMBL; AK[illegible]; BAB[illegible]
DR   MEROPS; [illegible]
DR   MGD; MGI:[illegible]
DR   InterPro; [illegible]
DR   InterPro; [illegible]
DR   InterPro; [illegible]
DR   Pfam; PF[illegible]IV_N[illegible]
DR   Pfam; [illegible]; Peptidase[illegible]
SQ   SEQUENCE   [illegible] AA; [illegible]

  Query Match          [illegible]
  Best Local Similarity   [illegible]
  Matches   [illegible]

Qy          1 ERHSIR 6
              [illegible]
Db        [illegible] ERHSIR [illegible]


RESULT 12
Q97GP1
ID   Q97GP1      PRELIMINARY;      [illegible] AA.
AC   Q97GP1;
DT   01-OCT-2001 (TrEMBLrel. [illegible])
DT   01-OCT-2001 (TrEMBLrel. [illegible] update)
DT   01-DEC-2001 (TrEMBLrel. [illegible] update)
DE   Nitroreductase family [illegible]
GN   CA[illegible]
OS   Clostridium acetobutylicum.
OC   Bacteria; Firmicutes; [illegible]
OC   Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1488;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 824 / DSM [illegible]
RX   MEDLINE=[illegible]
RA   Nolling J., [illegible] M., [illegible] K.,
RA   [illegible] M., [illegible] A. [illegible]
RA   [illegible]
RA   Bennett G.N., Koonin E.V., [illegible]
RT   "Genome sequence and [illegible]
RT   bacterium Clostridium acetobutylicum.";
RL   J. Bacteriol. [illegible]
DR   EMBL; AE[illegible]; AAK[illegible]
DR   InterPro; [illegible] Nitro[illegible]
DR   Pfam; PF[illegible]; Nitroreductase; [illegible]
KW   Complete proteome.
SQ   SEQUENCE   [illegible] AA; [illegible] CRC64;

  Query Match          [illegible]
  Best Local Similarity   [illegible]
  Matches   [illegible]

Qy          1 ERHSIR [illegible]
              [illegible]
Db        [illegible]


RESULT [illegible]
[illegible]
```

[illegible]

ID [illegible] PRELIMINARY; PRT; [illegible] AA.
DT [illegible] MAR [illegible] (TrEMBLrel. [illegible], Created)
DT [illegible] MAR [illegible] (TrEMBLrel. [illegible], Last sequence update)
DT [illegible] (TrEMBLrel. [illegible], Last annotation update)
DE [illegible] peptidase [illegible] (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE [illegible] PubMed [illegible]
RA Abbott C.A., Yu D.M.T., Woollatt E., Sutherland G.R., McCaughan G.W., [illegible]
RT "[illegible] expression and chromosomal localization [illegible]
RT dipeptidyl peptidase (DPP) IV homolog, DPP8.";
RL [illegible] Biochem. [illegible]
DR EMBL; AF [illegible]; AA [illegible]
DR MEROPS; S09.[illegible]
SQ SEQUENCE [illegible] AA; [illegible] MW; [illegible] CRC64;

Query Match [illegible] Score [illegible] DB [illegible] Length [illegible]
Best Local Similarity [illegible] Pred. No. [illegible]
Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

[illegible]

RESULT [illegible]
[illegible]
ID [illegible] PRELIMINARY; PRT; [illegible] AA.
AC [illegible]
DT 01 NOV 1998 (TrEMBLrel. 08, Created)
DT 01 NOV 1998 (TrEMBLrel. 08, Last sequence update)
DT [illegible] MAR [illegible] (TrEMBLrel. 20, Last annotation update)
DE [illegible] (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RA [illegible] Mc[illegible] P.M., [illegible] Adamson A.W., [illegible]
RA Burkhart-Schultz K., [illegible] Kyle A., [illegible]
RA [illegible]
RA [illegible]
RA [illegible] Attix C., Andreasen [illegible] Frankhouser M., [illegible]
RA [illegible]
RA [illegible]
RA [illegible] A.V.;
RT "[illegible]";
RL Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF [illegible]; AA [illegible]
DR [illegible]
DR [illegible]
DR [illegible]
DR [illegible] AMINOPEPTIDASE.
[illegible]
SQ SEQUENCE [illegible] AA; [illegible] MW; [illegible] CRC64;

Query Match [illegible] Score [illegible] DB [illegible] Length [illegible]
Best Local Similarity [illegible] Pred. No. [illegible]

Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

[illegible]

RESULT [illegible]
[illegible]
ID [illegible] PRELIMINARY; PRT; [illegible] AA.
AC [illegible]
DT 01 DEC [illegible] (TrEMBLrel. [illegible], Created)
DT 01 DEC [illegible] (TrEMBLrel. [illegible], Last sequence update)
DT 01 JUN [illegible] (TrEMBLrel. 21, Last annotation update)
DE cDNA FLJ[illegible] fis, clone PLACE1000[illegible], weakly similar to dipeptidyl
DE peptidase IV (EC 3.4.14.5).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Isogai T., [illegible] Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., [illegible] Sugano S., Shiratori A., Sudo H.,
RA Wagatsuma M., [illegible] Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., [illegible] Ishida S., Murakawa K., [illegible] Y., Takiguchi S.,
RA Watanabe S., [illegible] Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., [illegible]
RT "NEDO human cDNA sequencing project.";
RL Submitted (MAY [illegible]) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AK0[illegible]; BAA[illegible]
DR MEROPS; [illegible]
SQ SEQUENCE [illegible] AA; [illegible] MW; [illegible] CRC64;

Query Match [illegible] Score [illegible] DB [illegible] Length [illegible]
Best Local Similarity [illegible] Pred. No. [illegible]
Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

[illegible]

RESULT [illegible]
Q8WXI[illegible]
ID Q8WXI[illegible] PRELIMINARY; PRT; [illegible] AA.
AC Q8WXI[illegible];
DT 01 MAR [illegible] (TrEMBLrel. 20, Created)
DT 01 MAR [illegible] (TrEMBLrel. 20, Last sequence update)
DT 01 JUN [illegible] (TrEMBLrel. 21, Last annotation update)
DE Dipeptidyl peptidase [illegible].
GN DPP9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA [illegible]
RT "Identification and characterization of a novel member of the
RT dipeptidyl peptidase IV related family.";
RL Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF[illegible]; AA[illegible]
DR InterPro; [illegible] DPPIV_N_term.
DR InterPro; [illegible] Peptidase_S9.
DR InterPro; [illegible] aminopeptidase.
DR InterPro; [illegible] active site.
DR Pfam; PF[illegible] DPPIV_N_term; 1.
DR Pfam; PF[illegible] Peptidase_S9; 1.
DR PRINTS; [illegible] AMINOPTASE.
SQ SEQUENCE [illegible] AA; [illegible] MW; [illegible] CRC64;

[illegible]

```
;      STATE: [illegible]
;      COUNTRY: [illegible]
;      ZIP: [illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: [illegible]
;      COMPUTER: IBM Compatible
;      OPERATING SYSTEM: Windows
;      SOFTWARE: [illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible]
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME: [illegible]
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: [illegible]
;      TELEFAX: [illegible]
;      TELEX: [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: [illegible]
;      TYPE: nucleic acid
;      STRANDEDNESS: double
;      TOPOLOGY: [illegible]
;    MOLECULE TYPE: DNA [illegible]
;    HYPOTHETICAL: [illegible]
;    ANTI-SENSE: UNKNOWN
;    ORIGINAL SOURCE:
;      ORGANISM: [illegible]
;    FEATURE:
;      NAME/KEY: [illegible]
;      LOCATION: [illegible]
US [illegible]

Alignment Scores:
[illegible]
Score: [illegible]
Percent Similarity: [illegible]
Best Local Similarity: [illegible]
Query Match: [illegible]
DB: [illegible]

[illegible]

RESULT [illegible]
US [illegible]
; Sequence [illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;    APPLICANT: [illegible]
;    APPLICANT: [illegible]
;    APPLICANT: [illegible]
;    APPLICANT: [illegible]
;    APPLICANT: [illegible]
;    APPLICANT: [illegible]
```

[illegible]

[illegible]
MOLECULE TYPE: DNA (genomic)
[illegible]

[illegible]

RESULT 2
[illegible]
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MY[illegible]
; TITLE OF INVENTION: TUBERCULOSIS
[illegible]
; ORGANISM: Mycobacterium tuberculosis
[illegible]

[illegible]

RESULT [illegible]
[illegible]
; GENERAL INFORMATION:
[illegible]
; TITLE OF INVENTION: HUMAN [illegible]
; TITLE OF INVENTION: [illegible] AND DNA [illegible]
; NUMBER OF SEQUENCES: [illegible]

[illegible]

ALIGNMENTS

[illegible]

BASE COUNT [illegible]
ORIGIN

[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
Percent Similarity:	[illegible]	[illegible]	[illegible]
Best Local Similarity:	[illegible]	Mismatches:	[illegible]
Query Match:	100.00%	Indels:	[illegible]
DB:	[illegible]	Gaps:	[illegible]

[illegible]

[illegible]

```
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15q22"
                     /clone="121"
                     /tissue_type="placenta"
                     /clone_lib="lambdaTriplEx2[illegible]"
     gene            1..1147
                     /gene="DPP8"
     CDS             1..935
                     /gene="DPP8"
                     /note="serine protease; member of prolyl oligopeptidase
                     family belonging to enzyme clan SC; additional splice
                     variants of this mRNA are found in GenBank Accession
                     Numbers AF221634, AF221635 and AF221637"
                     /codon_start=1
                     /product="dipeptidyl peptidase 8"
                     /protein_id="AAG29767.1"
                     /db_xref="GI:11095190"
                     /translation="[illegible]
                     FISKYSNQKNP[illegible]
                     [illegible]
                     [illegible]
                     YLGSVAMQAEKFPSEPNFLLLHGF[illegible]ENVHFAHISILLSFLVKAGKPY[illegible]
                     HSIRVPESGEHYFLHLLHYLQEN[illegible]AALKVI"
BASE COUNT      357 a    258 c    242 g    341 t
ORIGIN

Alignment Scores:
Pred. No.:              341          Length:        1147
Score:                  31.00        Matches:       [illegible]
Percent Similarity:     100.00%      Conservative:  [illegible]
Best Local Similarity:  100.00%      Mismatches:    [illegible]
Query Match:            100.00%      Indels:        [illegible]
DB:                     9            Gaps:          [illegible]

SW-P3464-3 (1-6) x AF221635 (1-1147)

Qy          1 GluArgHisSerIleArg 6
              [illegible]
Db        825 [illegible] 842

RESULT 12
AF247777
LOCUS       AF247777              1253 bp    mRNA    linear   INV [illegible]
DEFINITION  Caenorhabditis elegans delta 12 fatty acid desaturase FAT-2 (fat-2)
            mRNA, complete cds.
ACCESSION   AF247777
VERSION     AF247777.1  GI:[illegible]
KEYWORDS    .
SOURCE      Caenorhabditis elegans.
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 1253)
  AUTHORS   Peyou-Ndi,M.M., Watts,J.L. and Browse,J.
  TITLE     Identification and Characterization of an Animal Delta 12 Fatty
            Acid Desaturase Gene by Heterologous Expression in Saccharomyces
            cerevisiae
  JOURNAL   Arch. Biochem. Biophys. 376 (2000) In press
REFERENCE   2  (bases 1 to 1253)
  AUTHORS   Peyou-Ndi,M.M., Watts,J.L. and Browse,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAR-2000) Institute of Biological Chemistry,
            Washington State University, Clark Hall, Pullman, WA [illegible]
            USA
FEATURES             Location/Qualifiers
     source          1..1253
                     /organism="Caenorhabditis elegans"
                     /db_xref="taxon:6239"
     gene            1..1253
                     /gene="fat-2"
     CDS             21..1151
                     /gene="fat-2"
```

```
[illegible]
BASE COUNT      [illegible]
ORIGIN

Alignment Scores:
Pred. No.:              [illegible]
Score:                  [illegible]
Percent Similarity:     [illegible]
Best Local Similarity:  [illegible]
Query Match:            [illegible]
DB:                     [illegible]

SW-P3464-3 [illegible]

Qy          [illegible]
Db          [illegible]

RESULT 13
AY0[illegible]
LOCUS       AY[illegible]
DEFINITION  Arabidopsis [illegible]
ACCESSION   AY[illegible]
VERSION     AY[illegible]
KEYWORDS    FLI-CDNA.
SOURCE      Arabidopsis thaliana
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; [illegible]
            Spermatophyta; Magnoliophyta; [illegible]
            [illegible]
REFERENCE   1  (bases 1 to [illegible])
  AUTHORS   [illegible]
  TITLE     Arabidopsis [illegible]
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to [illegible])
  AUTHORS   [illegible]
  TITLE     Direct Submission
  JOURNAL   Submitted [illegible]
COMMENT     [illegible]
```

```
                 [illegible]YAKNWEKI[illegible]YA[illegible]NIVVIK[illegible]VNA[illegible]NVV[illegible]A[illegible]IVLVAA
                 [illegible]AAY[illegible]A[illegible]NVIEKV[illegible]YA[illegible]EKAKV[illegible]KKEK[illegible]VIV[illegible]A[illegible]EKE[illegible]ALV[illegible]N[illegible]
                 [illegible]AP[illegible]HKVEL[illegible]"
BASE COUNT     [illegible]
ORIGIN

Alignment Scores:
Pred. No.:              [illegible]        Length:          [illegible]
Score:                  [illegible]        Matches:         [illegible]
Percent Similarity:     [illegible]        Conservative:    [illegible]
Best Local Similarity:  [illegible]        Mismatches:      [illegible]
Query Match:            [illegible]        Indels:          [illegible]
DB:                     [illegible]        Gaps:            [illegible]

SW[illegible] (1-[illegible]) x NEW[illegible]N (1-[illegible])

Qy      [illegible] [illegible]GAT[illegible]AT[illegible]A [illegible]
            [illegible]
Db      [illegible] [illegible]AAA[illegible]A[illegible]AATA[illegible]A [illegible]

RESULT [illegible]
AY[illegible]
LOCUS       AY[illegible]        [illegible] bp    mRNA    linear   PLN [illegible]-MAY-[illegible]
DEFINITION  Arabidopsis thaliana putative DNA binding protein (At[illegible]) mRNA,
            complete cds.
ACCESSION   AY[illegible]
VERSION     AY[illegible]  GI:[illegible]
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to [illegible])
  AUTHORS   Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,[illegible], [illegible],C.M.,
            [illegible],J.M., Goldsmith,A.D., Lee,J.M., [illegible], [illegible],H.L.,
            [illegible], [illegible],M., Wu,H.C., Yamamura,Y., Yu,[illegible], [illegible],
            [illegible], Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,[illegible],
            [illegible], Kamiya,A., Karlin-Neumann,[illegible], Kawai,[illegible], Kim,[illegible], [illegible],
            [illegible], Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., [illegible],
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Arabidopsis Open Reading Frame (ORF) Clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to [illegible])
  AUTHORS   Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,[illegible], [illegible],C.M.,
            [illegible],J.M., Goldsmith,A.D., Lee,J.M., [illegible], [illegible],H.L.,
            [illegible], [illegible],M., Wu,H.C., Yamamura,Y., Yu,[illegible], [illegible],
            [illegible], Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,[illegible],
            [illegible], Kamiya,A., Karlin-Neumann,[illegible], Kawai,[illegible], Kim,[illegible], [illegible],
            [illegible], Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., [illegible],
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted ([illegible]-APR-[illegible]) Plant Gene Expression Center, [illegible] Buchanan
            Street, Albany, CA [illegible], USA
COMMENT     The RIKEN Genomic Sciences Center ([illegible]) members contributed [illegible]
            [illegible] and [illegible] of RAFL cDNAs (RAFL cDNA : RIKEN
            Arabidopsis Full Length cDNA); Seki,M., Narusaka,M., [illegible],
            Satou,M., Kamiya,A., Sakurai,T., [illegible], Kawai,[illegible],
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) [illegible] members constructed and
            sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada,K.,
            Banh,J., Chan,M.M., Chang,C.H., Chang,[illegible], [illegible], [illegible],J.M.,
            Goldsmith,A.D., Lee,J.M., [illegible], [illegible],H.L., [illegible],
            [illegible],M., Wu,H.C., Yamamura,Y., Yu,[illegible], [illegible], Chen,H.,
            Cheuk,R., Jones,[illegible], Karlin-Neumann,[illegible], Kim,[illegible], [illegible],
            Meyers,M.C., Miranda,M., Nguyen,M., [illegible], Shinn,P.,
            Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

            Yamada,K. (SSP[illegible]) and Seki,M. (RIKEN [illegible]) contributed equally
            to this work. Shinozaki,K. (RIKEN [illegible]) and [illegible] (SSP
            [illegible]) contributed equally to this work as [illegible]
FEATURES             Location/Qualifiers
     source          1..[illegible]
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="1"
                     /clone="[illegible]"
                     /note="This clone is in pUNI [illegible]"
                     /ecotype="Columbia"
     gene            [illegible]
                     /gene="At1g[illegible]"
     CDS             [illegible]
                     /gene="At1g[illegible]"
                     /codon_start=1
                     /evidence=experimental
                     /product="putative DNA binding protein"
                     /protein_id="AAM[illegible].1"
                     /db_xref="GI:[illegible]"
                     /translation="M[illegible]KAMA[illegible]YYLQ[illegible]A[illegible]NLAL[illegible]
                     [illegible]NTHH[illegible]ANNP[illegible]HMGVVSSASLALVQ[illegible]LMVK[illegible]
                     [illegible]EMVSNKSRDSSPMS[illegible]NP[illegible]KRAR[illegible]DKQ[illegible]AN[illegible]WMN[illegible]
                     [illegible]V[illegible]GAG[illegible]AAKVLS[illegible]KRAA[illegible]MS[illegible]LSSVL[illegible]
                     [illegible]SYLVNL[illegible]SPSR[illegible]SVS[illegible]SRP[illegible]ST[illegible]VDML[illegible]AAN[illegible]VVA[illegible]
                     [illegible]HNNNNK[illegible]KFPNE[illegible]DNNSEMEL[illegible]SAALAASAA[illegible]
                     [illegible]SRSL[illegible]NEL[illegible]"
     3'UTR           [illegible]
                     /gene="At1g[illegible]"
BASE COUNT     [illegible]
ORIGIN

Alignment Scores:
Pred. No.:              [illegible]        Length:          [illegible]
Score:                  [illegible]        Matches:         [illegible]
Percent Similarity:     [illegible]        Conservative:    0
Best Local Similarity:  [illegible]        Mismatches:      0
Query Match:            [illegible]        Indels:          0
DB:                     [illegible]        Gaps:            0

SWOP3464 (1-[illegible]) x AY[illegible] (1-[illegible])

Qy      1 [illegible]GAT[illegible]AT[illegible] [illegible]
          [illegible]
Db    7[illegible] [illegible]A[illegible]AA[illegible] [illegible]

RESULT [illegible]
AF22[illegible]
LOCUS       AF[illegible]        [illegible] bp    mRNA    linear   [illegible]
DEFINITION  Homo sapiens dipeptidyl peptidase 8 (DPP8) mRNA, partial cds,
            alternatively spliced.
ACCESSION   AF[illegible]
VERSION     AF[illegible]  GI:[illegible]
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
            Mammalia; [illegible]; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases [illegible])
  AUTHORS   Abbott,C.A., Yu,D.M., Woollatt,E., Sutherland,G.R., McCaughan,G.W.
            and [illegible]
  TITLE     [illegible] expression and chromosomal localization of [illegible]
            dipeptidyl peptidase (DPP) IV homolog, DPP8
  JOURNAL   Eur. [illegible] 267 (20), [illegible] (2000)
  MEDLINE   [illegible]
   PUBMED   [illegible]
REFERENCE   2  (bases [illegible])
  AUTHORS   Abbott,C.A., Yu,D.M., McCaughan,G.W. and Gorrell,M.D.
  TITLE     Direct Submission
  JOURNAL   Submitted ([illegible]-JAN-2000) A.W. Morrow Gastroenterology [illegible]
            Centenary Institute of Cell Biology and Cancer [illegible]
            [illegible], Newtown, Sydney, NSW 2042, Australia
FEATURES             Location/Qualifiers
     source          1..[illegible]
                     /organism="Homo sapiens"
```

```
AX[illegible]
LOCUS       AX[illegible]                 [illegible] bp    NA     linear   [illegible]
DEFINITION  Sequence [illegible]
ACCESSION   AX[illegible]
VERSION     AX[illegible]
KEYWORDS    .
SOURCE      [illegible]
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Sharkey,R.A. and [illegible]M.
  TITLE     Nucleic acids [illegible]
            methods of use therefor
  JOURNAL   Patent: WO [illegible] MAY [illegible]
            [illegible]
FEATURES             Location/Qualifiers
     [illegible]
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     [illegible]
                     /note="Single nucleotide polymorphism
                     Accession number [illegible]"
BASE COUNT  [illegible]
ORIGIN

Alignment Scores:
Pred. No.:            [illegible]     Length:        51
Score:                [illegible]     Matches:       [illegible]
Percent Similarity:   [illegible]     Conservative:  [illegible]
Best Local Similarity: [illegible]    Mismatches:    [illegible]
Query Match:          [illegible]     Indels:        [illegible]
DB:                   [illegible]     Gaps:          [illegible]

US-[illegible] x AX[illegible]

Qy  [illegible]AAGHAS[illegible]

Db  [illegible]AAAA[illegible]

RESULT [illegible]
HUMNIT[illegible]
LOCUS       HUMNI[illegible]     [illegible] bp    DNA     linear   PRI [illegible]
DEFINITION  Human [illegible] endothelial nitric oxide synthase gene [illegible]
ACCESSION   [illegible]
VERSION     [illegible]
KEYWORDS    [illegible] synthase.
SEGMENT     [illegible]
SOURCE      Homo sapiens DNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases [illegible])
  AUTHORS   Marsden,P.A., Heng,H.H., Scherer,S.W., Stewart,R.J., Hall,A.V.,
            Shi,X.M., Tsui,L.C. and Schappert,K.
  TITLE     Structure and chromosomal localization of the human constitutive
            endothelial nitric oxide synthase gene
  JOURNAL   J. Biol. Chem. 268 (23), [illegible]
  MEDLINE   [illegible]
   PUBMED   [illegible]
FEATURES             Location/Qualifiers
     source          [illegible]
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     [illegible]     [illegible]
                     /number=22
     [illegible]     [illegible]
                     /number=2[illegible]
     [illegible]     [illegible]
                     /number=[illegible]
BASE COUNT  [illegible]
ORIGIN
```

[illegible]

ALIGNMENTS

[illegible]
at the University of [illegible]
[illegible]
Plasmid DNA preparations and DNA sequencing were
performed in the [illegible]

[illegible]

Alignment Scores
[illegible]

[illegible]

```
[illegible]
DEFINITION  [illegible] NCI_CGAP_[illegible] cDNA [illegible]
            similar to [illegible] SYNTHASE, [illegible] (HUMAN),
            mRNA sequence.
ACCESSION   [illegible]
VERSION     [illegible]
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to [illegible])
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     [illegible]
            [illegible]
            Unknown library type
            Insert Length: 401   Std Error: [illegible]
            Seq primer: [illegible]
FEATURES             Location/Qualifiers
     source          [illegible]
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:[illegible]"
                     /clone_lib="NCI_CGAP_[illegible]"
                     /sex="[illegible]"
                     /tissue_type="[illegible]"
                     /dev_stage="[illegible]"
                     /lab_host="DH10B"
                     /note="[illegible] cDNA [illegible]"
                     [illegible]
BASE COUNT   [illegible]
ORIGIN

[illegible]

SW:[illegible] x AI260[illegible] (1 [illegible])

[illegible]
DEFINITION  [illegible] NCI_CGAP_[illegible] cDNA [illegible]
            similar to [illegible] SYNTHASE, [illegible] (HUMAN),
```

SUMMARIES

ALIGNMENTS

```
PR   [illegible]-OCT-2000; 2000US-[illegible].
PR   [illegible]-OCT-2000; 2000US-[illegible].
PR   [illegible]-OCT-2000; 2000US-[illegible].
PR   [illegible]-OCT-2000; 2000US-[illegible].
PR   [illegible]-OCT-2000; 2000US-[illegible].
PR   [illegible]-OCT-2000; 2000US-[illegible].
PR   [illegible]-OCT-2000; 2000US-0241221.
PR   [illegible]-OCT-2000; 2000US-[illegible].
PR   [illegible]-OCT-2000; 2000US-[illegible].
PR   [illegible]-OCT-2000; 2000US-[illegible].
PR   [illegible]-OCT-2000; 2000US-0241808.
PR   [illegible]-OCT-2000; 2000US-[illegible].
PR   [illegible]-OCT-2000; 2000US-[illegible].
PR   [illegible]-NOV-2000; 2000US-[illegible].
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-[illegible].
PR   08-NOV-2000; 2000US-[illegible].
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-[illegible].
PR   08-NOV-2000; 2000US-[illegible].
PR   08-NOV-2000; 2000US-[illegible].
PR   08-NOV-2000; 2000US-[illegible].
PR   08-NOV-2000; 2000US-[illegible].
PR   08-NOV-2000; 2000US-[illegible].
PR   08-NOV-2000; 2000US-[illegible].
PR   08-NOV-2000; 2000US-[illegible].
PR   08-NOV-2000; 2000US-[illegible].
PR   08-NOV-2000; 2000US-[illegible].
PR   08-NOV-2000; 2000US-[illegible].
PR   08-NOV-2000; 2000US-[illegible].
PR   17-NOV-2000; 2000US-[illegible].
PR   17-NOV-2000; 2000US-[illegible].
PR   17-NOV-2000; 2000US-[illegible].
PR   17-NOV-2000; 2000US-[illegible].
PR   17-NOV-2000; 2000US-[illegible].
PR   17-NOV-2000; 2000US-[illegible].
PR   17-NOV-2000; 2000US-[illegible].
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-[illegible].
PR   17-NOV-2000; 2000US-[illegible].
PR   17-NOV-2000; 2000US-[illegible].
PR   17-NOV-2000; 2000US-[illegible].
PR   17-NOV-2000; 2000US-[illegible].
PR   17-NOV-2000; 2000US-[illegible].
PR   17-NOV-2000; 2000US-[illegible].
PR   17-NOV-2000; 2000US-[illegible].
PR   17-NOV-2000; 2000US-[illegible].
PR   17-NOV-2000; 2000US-[illegible].
PR   17-NOV-2000; 2000US-[illegible].
PR   [illegible]-DEC-2000; 2000US-[illegible].
PR   [illegible]-DEC-2000; 2000US-[illegible].
PR   05-DEC-2000; 2000US-[illegible].
PR   05-DEC-2000; 2000US-[illegible].
PR   05-DEC-2000; 2000US-[illegible].
PR   06-DEC-2000; 2000US-[illegible].
PR   08-DEC-2000; 2000US-[illegible].
PR   08-DEC-2000; 2000US-[illegible].
PR   08-DEC-2000; 2000US-[illegible].
PR   08-DEC-2000; 2000US-[illegible].
PR   08-DEC-2000; 2000US-[illegible].
PR   11-DEC-2000; 2000US-[illegible].
PR   05-JAN-2001; 2001US-[illegible].
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Barash SC, Ruben SM;
XX
DR   WPI; 2001-[illegible]/50.
DR   P-PSDB; AAU[illegible].
XX
PT   Novel polypeptides and polynucleotides useful for diagnosing or
PT   preventing, treating neural, immune system, muscular, reproductive,
PT   pulmonary, cardiovascular, renal, proliferative disorders and [illegible]
PT   diseases
XX
PS   Claim [illegible]
XX
CC   The present invention [illegible]
CC   polypeptides [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   diagnosis, treatment [illegible]
CC   disorders [illegible]
CC   immunodeficiency [illegible]
CC   (e.g. arthritis) [illegible]
CC   metabolic disorders [illegible]
CC   (e.g. asthma) [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   Note: [illegible]
CC   [illegible]
CC   [illegible]
XX
SQ   Sequence [illegible]

Alignment Scores:
Pred. No.: [illegible]
Score: [illegible]
Percent Similarity: [illegible]
Best Local Similarity: [illegible]
Query Match: [illegible]
DB: [illegible]

[illegible]

[illegible]

RESULT [illegible]
AA[illegible]
ID   AA[illegible] standard; [illegible]
XX
AC   AA[illegible]
XX
DT   [illegible]
XX
DE   [illegible]
XX
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
XX
OS   [illegible]
XX
FH   [illegible]
FT   [illegible]
FT   [illegible]
FT   [illegible]
XX
PN   W[illegible]
XX
PD   [illegible]
XX
PF   [illegible]
XX
PR   [illegible]
PR   [illegible]
XX
PA   [illegible]
```

```
[illegible]
XX
[illegible]
XX
PT   Novel [illegible] isolated polypeptides, useful for diagnosis, treatment and
PT   prevention of neural, immune system, muscular, reproductive,
PT   gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT   disorders.
XX
[illegible]
XX
CC   The invention relates to novel genes ([illegible]) and proteins
CC   ([illegible]) useful for preventing, treating or ameliorating
CC   medical conditions, e.g. by protein or gene therapy. The genes are
CC   isolated from a number of human tissues disclosed in the specification.
CC   The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC   in the diagnosis, treatment and prevention of: (1) cancer, e.g. breast
CC   [illegible]
CC   [illegible]
CC   [illegible] Addison's disease [illegible]
CC   [illegible]
CC   [illegible] multiple sclerosis [illegible]
CC   [illegible] (3) cardiovascular disorders [illegible]
CC   [illegible]
CC   [illegible] (5) infectious diseases such as viral, bacterial, fungal
CC   and parasitic infections.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
[illegible]
```

[illegible]

```
CC   proteins and polynucleotides may be used [illegible]
CC   treatment of diseases associated with inappropriate (I) expression [illegible]
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's [illegible]
CC   that affect the activity of (I) or expression [illegible]
CC   supplement the patient's own production [illegible]
CC   [illegible]
CC   [illegible]
CC   protein. (I) proteins and polynucleotides can be used to prevent,
CC   diagnose and treat immune/haematopoietic related diseases, especially
CC   cancers and cancer metastasis of haematopoietic derived cells. AAK[illegible]
CC   to AAK[illegible] represent human immune/haematopoietic [illegible]
CC   sequences of the present invention. AAK[illegible] AAK[illegible] and AAK[illegible]
CC   represent sequences used in the exemplification of the present [illegible]
XX
SQ   Sequence 688 BP; 208 A; 1[illegible] C; 1[illegible] G; [illegible] other;

Alignment Scores:
[illegible]
Percent Similarity:  [illegible]     [illegible]
Best Local Similarity:  100.00%     Mismatches: [illegible]
Query Match:  100.0[illegible]     Indels: [illegible]
DB:  22     Gaps: [illegible]

SW[illegible] x AAK[illegible]

Qy       1 [illegible]ArgHisSerIleArg[illegible]
           [illegible]
Db      1[illegible] [illegible]AAA[illegible]A[illegible]

[illegible]
AAK[illegible]
ID   AAK[illegible] standard; cDNA; [illegible] BP.
XX
AC   AAK[illegible]
XX
DT   [illegible] (first entry)
XX
DE   Human protease and protease inhibitor [illegible] cDNA
XX
KW   Human protease; protease inhibitor; [illegible] and protease inhibitor [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible] immune suppressive; [illegible]
KW   [illegible] antibacterial; [illegible]
KW   antiatherosclerotic; antiplatelet; [illegible]
KW   autoimmune disorder; inflammatory disease; [illegible]
KW   [illegible]
KW   [illegible] disease; AIDS; [illegible] disease; [illegible]
KW   [illegible] disease; diabetes; [illegible]
KW   [illegible] disease; Hashimoto's thyroiditis; [illegible]
KW   Werner's syndrome; cell proliferative disorder; arterial [illegible]
KW   atherosclerosis; cirrhosis; hepatitis; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   W[illegible]-A[illegible]
XX
PD   [illegible]-2001.
XX
PF   [illegible] 2000; [illegible] W[illegible]
XX
PR   [illegible]
PR   [illegible]
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   [illegible]
PI   Yue H;
XX
DR   WPI; [illegible]/[illegible]
DR   P-PSDB; AAB[illegible]
```

```
XX
PT   New [illegible]
PT   [illegible]
PT   [illegible]
PT   [illegible]
XX
PS   [illegible]
XX
CC   [illegible]
CC   (PRIMP) [illegible]
CC   [illegible]
CC   [illegible]
XX
SQ   Sequence [illegible]

Alignment Scores:
[illegible]

SW[illegible]

Qy  [illegible]

Db  [illegible]

[illegible]
AAK[illegible]
ID   AAK[illegible] standard; cDNA; [illegible]
XX
AC   [illegible]
XX
DT   [illegible]
XX
DE   Human p[illegible]
XX
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
XX
OS   [illegible]
XX
PN   W[illegible]
XX
PD   [illegible]
XX
PF   [illegible]
XX
PR   [illegible]
XX
PA   (INCY-) [illegible]
XX
```

```
XX
PN   EP1[illegible]04-A2.
XX
PD   0[illegible]-SEP-2001.
XX
PF   [illegible]-JUL-2000; 2000EP-[illegible].
XX
PR   [illegible]-JUL-1999;   99JP-0194186.
PR   [illegible]-JAN-2000; 2000JP-0118771.
PR   [illegible]-MAY-2000; 2000JP-0183765.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Nishikawa T,  Isogai T,  Hayashi K,  [illegible],  Kawai Y;
PI   Wakamatsu A,  Sugiyama T,  Nagai K,  Fujimori S,  Otsuki T,  Kato H;
XX
DR   WPI; 2001-524255/58.
XX
PT   New primers useful for synthesizing full length cDNA clones and their
PT   use in genetic manipulation -
XX
PS   Claim 2; SEQ ID NO 543; 1580pp + sequence listing; English.
XX
CC   The invention relates to primers for synthesising full length cDNA
CC   clones. 830 cDNA molecules encoding a human protein have been
CC   isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC   molecules have been determined. Primers for synthesising the full length
CC   cDNA are useful for clarifying the function of the protein encoded by
CC   the cDNA. The full length clones were obtained by construction of full
CC   length enriched cDNA libraries that were synthesised by the oligo-capping
CC   method. The primers enable the production of the full length cDNA easily
CC   without any special methods. The present sequence is the nucleotide
CC   sequence of the 5'-end of a cDNA provided in the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in CD-ROM format directly from EPO
XX
SQ   Sequence 662 BP; 159 A; 20[illegible] C; 182 G; 115 [illegible] other;

Alignment Scores:
Pred. No.:              557            Length:        [illegible]
Score:                  [illegible]    Matches:       [illegible]
Percent Similarity:     100.00%        Conservative:  [illegible]
Best Local Similarity:  100.00%        Mismatches:    [illegible]
Query Match:            100.00%        Indels:        [illegible]
DB:                     22             Gaps:          [illegible]

SW [illegible] (1-6) x AAK[illegible] ([illegible])

Qy        [illegible] ArgIleSerIleArg [illegible]
          [illegible]
Db    [illegible] [illegible]AGA[illegible]A[illegible]AT[illegible] [illegible]

RESULT [illegible]
AAK[illegible]
ID   AAK[illegible] standard; cDNA; [illegible] BP.
XX
AC   AAK[illegible];
XX
DT   [illegible]-NOV-2001 (first entry)
XX
DE   Human cDNA clone representative sequence, SEQ ID NO: 1[illegible].
XX
KW   Human; full length cDNA; cDNA synthesis; [illegible]; ss.
XX
OS   Homo sapiens.
XX
PN   EP1[illegible]04-A2.
XX
PD   05-SEP-2001.
XX
PF   07-JUL-2000; 2000EP-[illegible].
XX
PR   08-JUL-1999;   99JP-0194186.
```

```
PR   [illegible]-JAN-[illegible]
PR   [illegible]-MAY-[illegible]
XX
PA   (HELI-) [illegible] RES INST.
XX
PI   [illegible] N[illegible]
PI   Wakamatsu A, [illegible]
XX
DR   WPI; [illegible]
XX
PT   [illegible] NA [illegible]
PT   use in genetic manipulation [illegible]
XX
PS   Example [illegible] SEQ [illegible]
XX
CC   [illegible]
CC   clones. 8[illegible] cDNA [illegible]
CC   isolated [illegible]
CC   molecules have been determined [illegible]
CC   cDNA are useful for clarifying [illegible]
CC   the cDNA. [illegible]
CC   length enriched cDNA libraries [illegible]
CC   method. The primers enable [illegible]
CC   without any special methods [illegible]
CC   representative [illegible]
CC   human [illegible]
CC   Note: The sequence data [illegible]
CC   specification [illegible]
XX
SQ   Sequence [illegible] BP; [illegible] A; [illegible]

Alignment Scores:
Pred. No.:              [illegible]
Score:                  [illegible]
Percent Similarity:     [illegible]
Best Local Similarity:  [illegible]
Query Match:            [illegible]
DB:                     [illegible]

SW [illegible] ([illegible]) x AAK[illegible] ([illegible])

Qy        [illegible] A [illegible]
          [illegible]
Db    [illegible] [illegible]AGA[illegible]A[illegible]

RESULT [illegible]
AAK[illegible]
ID   AAK[illegible] standard; [illegible]
XX
AC   AAK[illegible];
XX
DT   [illegible]-NOV-[illegible] (first entry)
XX
DE   Human [illegible]
XX
KW   Human; [illegible]
KW   [illegible]
XX
OS   [illegible]
XX
PN   W[illegible] A[illegible]
XX
PD   [illegible] A [illegible]
XX
PF   [illegible]-JAN-[illegible] W[illegible]
XX
PR   [illegible]-JAN-[illegible]
PR   [illegible]-FEB-[illegible]
PR   [illegible]-FEB-[illegible]
PR   [illegible]-MAR-[illegible]
PR   [illegible]-MAR-[illegible]
PR   [illegible]-MAR-[illegible]
PR   [illegible]-A[illegible]
```

```
XX
PA   [illegible]
XX
PI   [illegible]
XX
DR   [illegible]
XX
PT   [illegible]
PT   cells useful for the prevention, diagnosis and treatment of [illegible]
PT   infestations
XX
PS   Claim [illegible]; Page [illegible]; [illegible]pp; English.
XX
CC   The invention relates to novel cat flea (Ctenocephalides felis) [illegible]
CC   acids which are expressed in hindgut and Malpighian tubule (HMT) [illegible]
CC   [illegible] hindgut and nerve cord (HNC) tissues. The invention also relates to the
CC   encoded proteins. The invention additionally [illegible] expression
CC   constructs, recombinant viruses and recombinant [illegible]
CC   nucleic acids of the invention, recombinant [illegible]
CC   antibodies against the proteins, [illegible]
CC   the proteins and [illegible]
CC   [illegible]
CC   [illegible] may be used in the prevention, treatment [illegible] diagnosis [illegible]
CC   associated with flea infestations. [illegible]
CC   used to produce an HMT or HNC protein [illegible]
CC   DNA [illegible] by inserting the nucleic [illegible]
CC   [illegible] the host to express the protein. The HMT and HNC [illegible]
CC   may also be used as DNA probes in diagnostic assays [illegible]
CC   and quantitate the presence of [illegible]
CC   sequences in samples. They may also [illegible]
CC   [illegible] of the proteins and their [illegible] The HMT and HNC
CC   proteins may be used as antigens in the production of specific
CC   antibodies, and in assays to identify [illegible] (agonists and
CC   antagonists) of HMT and/or HNC protein expression and activity. [illegible]
CC   anti HMT/HNC protein antibodies and antagonists may also be used [illegible]
CC   downregulate protein expression and activity. The antibodies may also be
CC   used as diagnostic agents for detecting the presence of flea polypeptides
CC   in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The
CC   present sequence represents a cat flea HMT cDNA of the invention.
XX
SQ   Sequence [illegible]

Alignment Scores:
[illegible]

Sw [illegible] AA [illegible]

Qy   [illegible]

Db   [illegible] AA AA [illegible]

RESULT [illegible]
AW[illegible]
ID   AW[illegible] standard; cDNA; [illegible]
XX
AC   AW[illegible];
XX
DT   [illegible] (first entry)
XX
DE   Human prostate expression marker cDNA [illegible]
XX
KW   Human prostate [illegible]
KW   [illegible] marker; [illegible]
XX
OS   [illegible]
XX
PN   W[illegible]
XX

[illegible]
XX
[illegible]
XX
PA   [illegible]
XX
PI   [illegible]
XX
DR   W[illegible]
XX
PT   [illegible]
XX
PS   [illegible]
XX
CC   [illegible]
XX
SQ   [illegible]

Alignment Scores:
[illegible]

RESULT [illegible]
ID   [illegible]
XX
AC   [illegible]
XX
DT   [illegible]
XX
DE   [illegible]
XX
KW   [illegible]
XX
OS   [illegible]
XX
PN   [illegible]
XX
PD   [illegible]
XX
```

[illegible]
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI [illegible] Monahan [illegible]
XX
DR WPI; [illegible]
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer
XX
PS Claim [illegible]; [illegible]pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in tables [illegible] (AbV[illegible] AbV[illegible]) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence [illegible]

Alignment Scores:
Pred. No.: [illegible] Length: [illegible]
Score: [illegible] Matches: [illegible]
Percent Similarity: [illegible] Conservative: [illegible]
Best Local Similarity: [illegible] Mismatches: [illegible]
Query Match: [illegible] Indels: [illegible]
DB: [illegible] Gaps: [illegible]

[illegible]

[illegible]

XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer
XX
PS Claim [illegible]; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in tables [illegible] (ABV[illegible] ABV[illegible]) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence [illegible]

Alignment Scores:
Pred. No.: [illegible] Length: 416
Score: [illegible] Matches: [illegible]
Percent Similarity: [illegible] Conservative: [illegible]
Best Local Similarity: [illegible] Mismatches: [illegible]
Query Match: [illegible] Indels: [illegible]
DB: [illegible] Gaps: [illegible]

[illegible]

XX
KW [illegible] nucleic acid; BMI.
KW [illegible] therapeutic target;
KW [illegible]
XX
FN [illegible]
XX
[illegible]

```
XX
PD   [illegible]-MAY-2001.
XX
PF   22-NOV-2000; 2000WO-US[illegible].
XX
PR   24-NOV-1999;   99US-0167384.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA, [illegible] M;
XX
DR   WPI; 2001-[illegible]/[illegible].
XX
PT   Isolated human nucleic acids comprising one or more single nucleotide
PT   polymorphisms, useful for treating a subject suffering from a
PT   pathology, e.g. autoimmune diseases, ascribed to the presence of a
PT   sequence polymorphism
XX
PS   Claim 1; Page 441; 674pp; English.
XX
CC   AAH[illegible] to AA[illegible]104 represent human nucleic acid oligonucleotides
CC   comprising one or more single nucleotide polymorphisms (SNPs). AAB[illegible]
CC   to AAB[illegible] represent human peptides encoded by some of the SNP
CC   oligonucleotides. The sequences from the present invention can have
CC   immunosuppressive, cytostatic, antiinflammatory, neuroprotective and
CC   antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides
CC   and antibodies from the present invention can be used for treating a
CC   subject suffering from, at risk for, or suspected of, suffering from a
CC   pathology ascribed to the presence of a sequence polymorphism. The
CC   pathology may be autoimmune diseases, inflammation, cancer, diseases of
CC   the nervous system, and infection by pathogenic microorganisms. The SNPs
CC   are also useful for determining which forms of a characterised
CC   polymorphism are present in individuals. The antibodies may be used in
CC   the detection, quantitation and/or cellular or tissue localisation of
CC   polymorphic protein (e.g., for use in measuring levels of the
CC   polymorphic protein within appropriate physiological samples).
XX
SQ   Sequence 51 BP; 18 A; 10 C; 1[illegible] G; 1[illegible] T; [illegible] other;

Alignment Scores:
Pred. No.:              29.[illegible]        Length:        51
Score:                  31.00        Matches:       [illegible]
Percent Similarity:     100.00%      Conservative:
Best Local Similarity:  100.00%      Mismatches:
Query Match:            100.00%      Indels:
DB:                     2[illegible]            Gaps:

SW[illegible] (1-6) x AAB[illegible] (1-51)

Qy        1 [illegible]GluArgHisSerIleArg 6
            [illegible]
Db       [illegible] [illegible]AGAAT[illegible]AAG[illegible] 2[illegible]

RESULT 2
ABV[illegible]
ID   ABV[illegible] standard; cDNA; [illegible] BP.
XX
AC   ABV[illegible];
XX
DT   1[illegible]-SEP-2002 (first entry)
XX
DE   Human prostate expression marker cDNA [illegible].
XX
KW   Human; prostate cancer; cytostatic; [illegible] pharmacodynamic [illegible]
KW   pharmacogenomic marker; [illegible].
XX
OS   Homo sapiens.
XX
PN   WO200[illegible] A2.
XX
PD   2[illegible]-AUG-2001.
XX
PF   2[illegible]-FEB-2001; 2001WO-US0[illegible].
```

```
XX
PR   [illegible]-FEB-[illegible]   [illegible]S[illegible]
PR   [illegible]-MAR-[illegible]   [illegible]S[illegible]
PR   [illegible]-MAY-[illegible]   [illegible]S[illegible]
PR   [illegible]-JUN-[illegible]   [illegible]S[illegible]
PR   [illegible]-JUL-[illegible]   [illegible]S[illegible]
PR   [illegible]-[illegible]   [illegible]S[illegible]
XX
PA   ([illegible]) M[illegible]
XX
PI   [illegible]
XX
DR   WPI; [illegible].
XX
PT   [illegible]
PT   prostate [illegible]
PT   [illegible]
XX
PS   Claim [illegible]
XX
CC   The invention [illegible]
CC   [illegible] sequence [illegible] AB[illegible]
CC   specification [illegible]
CC   (a) assessing whether [illegible] with prostate [illegible]
CC   (b) monitoring the [illegible]
CC   (c) assessing the [illegible]
CC   cancer in a patient [illegible]
CC   (d) assessing [illegible]
CC   in a patient [illegible]
CC   (e) selecting [illegible]
CC   (f) assessing the prostate [illegible]
CC   (g) determining whether [illegible]
CC   (h) assessing the [illegible] prostate [illegible]
CC   patient [illegible]
CC   (i) [illegible]
XX
SQ   Sequence [illegible] A [illegible]

Alignment Scores:
Pred. No.:              [illegible]
Score:                  [illegible]
Percent Similarity:     [illegible]
Best Local Similarity:  [illegible]
Query Match:            [illegible]
DB:                     [illegible]

SW[illegible] ([illegible]) x ABV[illegible] ([illegible])

Qy        1 [illegible]
            [illegible]
Db       [illegible] [illegible]AAA[illegible]

RESULT [illegible]
ABV[illegible]
ID   ABV[illegible] standard; cDNA; [illegible]
XX
AC   ABV[illegible];
XX
DT   [illegible] (first entry)
XX
DE   Human prostate [illegible]
XX
KW   Human; prostate cancer; [illegible]
KW   pharmaco[illegible] marker; [illegible]
XX
OS   Homo sapiens.
XX
PN   W[illegible] A[illegible].
XX
PD   [illegible]-AUG-[illegible].
XX
PF   [illegible]
XX
```

```
;           TELECOMMUNICATION INFORMATION:
;                TELEPHONE: (415) [illegible]
;                TELEFAX: (415) [illegible]
;     INFORMATION FOR SEQ ID NO: 16:
;           SEQUENCE CHARACTERISTICS:
;                LENGTH: 1284 base pairs
;                TYPE: nucleic acid
;                STRANDEDNESS: single
;                TOPOLOGY: linear
;          MOLECULE TYPE: DNA (genomic)
;          SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-[illegible]

Alignment Scores:
Pred. No.:                  141              Length:            1284
Score:                      [illegible]      Matches:           [illegible]
Percent Similarity:         77.78%           Conservative:      1
Best Local Similarity:      66.67%           Mismatches:        2
Query Match:                [illegible]      Indels:
DB:                         4                Gaps:

SW [illegible] (1-9) x US-[illegible] (1-1284)

Qy          1 LeuAspGluAsnValHisPheAlaHis 9
DB       [illegible] [illegible]GA[illegible]CA[illegible]CA[illegible]AC 1[illegible]

RESULT 1[illegible]
US-[illegible]-444-1/c
; Sequence 1, Application US/09[illegible]444
; Patent No. 625[illegible]
; GENERAL INFORMATION:
;  APPLICANT: Cook, James W.
;  APPLICANT: Bulawa, Christine
;  TITLE OF INVENTION: ASPERGILLUS NIDULANS [illegible] TRANSFERASE GENES AND [illegible]
;  TITLE OF INVENTION: AND USES THEREOF
;  FILE REFERENCE: 07334/089001
;  CURRENT APPLICATION NUMBER: US/09/163,444
;  CURRENT FILING DATE: 1998-[illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: FastSEQ for Windows Version [illegible]
; SEQ ID NO [illegible]
;   LENGTH: 1[illegible]
;   TYPE: DNA
;   ORGANISM: Aspergillus fumigatus [illegible]
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...([illegible])
US-[illegible]-444-1

Alignment Scores:
Pred. No.:                  155              Length:            [illegible]
Score:                      [illegible]      Matches:           [illegible]
Percent Similarity:         [illegible]      Conservative:      [illegible]
Best Local Similarity:      [illegible]      Mismatches:        [illegible]
Query Match:                [illegible]      Indels:
DB:                         4                Gaps:

SW [illegible] (1-9) x US-[illegible]-444-1 (1-[illegible])

Qy          4 AsnValHisPheAlaHis [illegible]
Db       [illegible] AA[illegible]CA[illegible]GCGCA[illegible]

RESULT 14
US-[illegible]-444-2/c
; Sequence 2, Application US/[illegible]444
; Patent No. 62[illegible]
; GENERAL INFORMATION:
;  APPLICANT: Cook, James W.
;  APPLICANT: Bulawa, Christine
;  TITLE OF INVENTION: ASPERGILLUS NIDULANS [illegible] TRANSFERASE GENES AND [illegible]
;  TITLE OF INVENTION: AND USES THEREOF
```

```
[illegible]
; CURRENT APPLICATION NUMBER: [illegible]
; CURRENT FILING DATE: [illegible]
; NUMBER OF SEQ ID NOS: [illegible]
; SOFTWARE: [illegible]
; SEQ ID NO [illegible]
; [illegible]
;   TYPE: DNA
;   ORGANISM: [illegible]
US-[illegible]-111

Alignment Scores:
Pred. No.:                  [illegible]
Score:                      [illegible]
Percent Similarity:         [illegible]
Best Local Similarity:      [illegible]
Query Match:                [illegible]
DB:                         4

SW [illegible]

Qy          [illegible]
Db       [illegible] AA[illegible]

RESULT 1[illegible]
US-[illegible]
; Sequence [illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  TITLE OF INVENTION: [illegible]
;  TITLE OF INVENTION: [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: [illegible]
;  CURRENT FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: [illegible]
; SEQ ID NO [illegible]
;   LENGTH: [illegible]
;   TYPE: DNA
;   ORGANISM: [illegible]
US-[illegible]

Alignment Scores:
Pred. No.:                  [illegible]
Score:                      [illegible]
Percent Similarity:         [illegible]
Best Local Similarity:      [illegible]
Query Match:                [illegible]
DB:                         [illegible]

SW [illegible]

Qy          [illegible]

Db       [illegible]

Search completed: [illegible]
Job time : [illegible]
```

NAME/KEY:

GENERAL INFORMATION

Lueng Matches

SW-REPT

GENERAL INFORMATION

APPLICANT:

APPLICANT:

APPLICANT:

APPLICANT:

TITLE OF INVENTION:

TITLE OF INVENTION:

FILE REFERENCE:

CURRENT APPLICATION NUMBER:

CURRENT FILING DATE:

NUMBER OF SEQ ID NOS:

SOFTWARE:

SEQ ID NO

LENGTH: 11

TYPE: DNA

ORGANISM:

Alignment Scores:

Pred. No.:

Length:

Score:

Matches:

Percent Similarity:

Conservative:

Best Local Similarity:

Mismatches:

Query Match:

Indels:

DB:

Gaps:

GENERAL INFORMATION

APPLICANT:

TITLE OF INVENTION:

NUMBER OF SEQ ID NOS:

COMPUTER READABLE FORM:

CURRENT APPLICATION DATA:

PRIOR APPLICATION DATA:

ATTORNEY/AGENT INFORMATION:

NAME:

REGISTRATION NUMBER:

REFERENCE/DOCKET NUMBER:

```
;    NUMBER OF SEQUENCES: 24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Scully, Scott, Murphy & Presser
;      STREET: 400 Garden City Plaza
;      CITY: Garden City
;      STATE: New York
;      COUNTRY: U.S.A.
;      ZIP: 11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/502,046
;      FILING DATE: 14-JUL-1995
;      CLASSIFICATION: [illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 08/285,[illegible]
;      FILING DATE: [illegible]-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME: [illegible], Frank S.
;      REGISTRATION NUMBER: 31,[illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (516) 742-4343
;      TELEFAX: (516) 742-4366
;      TELEX: [illegible] SANS UR
;  INFORMATION FOR SEQ ID NO: [illegible]:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 1872 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: single
;      TOPOLOGY: linear
;    MOLECULE TYPE: DNA (genomic)
US-08-502-046-2[illegible]

Alignment Scores:
Pred. No.:              152             Length:         [illegible]
Score:                  37.00           Matches:        [illegible]
Percent Similarity:     [illegible]     Conservative:   [illegible]
Best Local Similarity:  [illegible]     Mismatches:     [illegible]
Query Match:            72.55%          Indels:         [illegible]
DB:                     2               Gaps:           [illegible]

SW[illegible] [illegible]

Qy       [illegible]AspI[illegible]AsnValHisPheAlaHis[illegible]
                [illegible]
Db       [illegible]AAAA[illegible]

RESULT 8
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. 61[illegible]
; GENERAL INFORMATION:
;    APPLICANT: FUKUCHI, Yasuhiro
;    APPLICANT: F[illegible]YKAWA, Sh[illegible]
;    APPLICANT: SHIMADA, Yukihi[illegible]
;    APPLICANT: [illegible]BAYASHI, Masaya
;    APPLICANT: SHIMADA, Hi[illegible]
;    APPLICANT: [illegible]INAKA, Yasu[illegible]
;    TITLE OF INVENTION: NOVEL PLANT GENE
;    NUMBER OF SEQUENCES: 67
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
;      STREET: [illegible] Rockefeller Plaza
;      CITY: New York
;      STATE: New York
;      COUNTRY: U.S.A.
;      ZIP: [illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Diskette, 3.50 inch, [illegible]
```

[illegible]

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: [illegible]
;    COMPUTER READABLE FORM:
;      NAME: [illegible]
;      TELEFAX: [illegible]
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: [illegible]
;      TYPE: [illegible]
;    FEATURE:
;      NAME/KEY: [illegible]

Alignment Scores:
[illegible]

RESULT [illegible]
US-[illegible]
; Sequence [illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;    APPLICANT: [illegible]
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      CITY: New York
;      STATE: New York
;      COUNTRY: U.S.A.
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: [illegible]
;      COMPUTER: IBM PC [illegible]
[illegible]
```

```
; Sequence [illegible], Application [illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;   APPLICANT: [illegible]
;   TITLE OF INVENTION: Streptococcus [illegible]
;   NUMBER OF SEQUENCES: [illegible]
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Human Genome Sciences, Inc.
;     STREET: [illegible] Key West Avenue
;     CITY: Rockville
;     STATE: Maryland
;     COUNTRY: USA
;     ZIP: [illegible]
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette, [illegible]
;     COMPUTER: [illegible]
;     OPERATING SYSTEM: MS-DOS [illegible]
;     SOFTWARE: ASCII Text
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: [illegible]
;     FILING DATE:
;     CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Brookes, A. Anders
;     REGISTRATION NUMBER: [illegible]
;     REFERENCE/DOCKET NUMBER: [illegible]
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: [illegible]
;     TELEFAX: [illegible]
; INFORMATION FOR SEQ ID NO: [illegible]:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: [illegible] base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
US-[illegible]

Alignment Scores:
Pred. No.:              [illegible]      Length:        [illegible]
Score:                  [illegible]      Matches:       [illegible]
Percent Similarity:     [illegible]      Conservative:  [illegible]
Best Local Similarity:  [illegible]      Mismatches:    [illegible]
Query Match:            76.47%           Indels:        [illegible]
DB:                     4                Gaps:          [illegible]

SW[illegible]

[illegible] AspGlnAsnValHisGluAlaHis [illegible]

[illegible] AAAATATACA [illegible]

RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application [illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;   APPLICANT: [illegible]
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING [illegible]
;   TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;   FILE REFERENCE: [illegible]
;   CURRENT APPLICATION NUMBER: [illegible]
;   CURRENT FILING DATE: [illegible]
;   PRIOR APPLICATION NUMBER: US 60/064,964
;   PRIOR FILING DATE: 1997-11-[illegible]
;   PRIOR APPLICATION NUMBER: [illegible]
;   PRIOR FILING DATE: 1997-[illegible]
;   NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO [illegible]
;   LENGTH: [illegible]
;   TYPE: DNA
```

[illegible]

SUMMARIES

ALIGNMENTS

; CURRENT APPLICATION NUMBER: US/09/200,114A

SWOPE464

```
                    /rpt_family "GT rich"
     repeat_region  12008..12028
                    /rpt_family "(GT)n"
     repeat_region  12084..12161
                    /rpt_family "A-rich"
     repeat_region  complement(12722..12877)
                    /rpt_family "MIR"
     repeat_region  13023..13086
                    /rpt_family "L3"
     repeat_region  13212..13282
                    /rpt_family "L2"
     repeat_region  13283..13537
                    /rpt_family "L2"
     repeat_region  13538..13667
                    /rpt_family "FLAM_C"
     repeat_region  13668..14227
                    /rpt_family "L2"
     repeat_region  complement(14228..14513)
                    /rpt_family "AluSx"
     repeat_region  14514..14525
                    /rpt_family "L2"
     repeat_region  complement(14526..14835)
                    /rpt_family "AluSx"
     repeat_region  14836..14999
                    /rpt_family="L2"
     repeat_region  complement(15356..15687)
                    /rpt_family="L2"
     repeat_region  15859..15991
                    /rpt_family="MIR"
     repeat_region  16064..16363
                    /rpt_family "AluSq"
     repeat_region  16597..16809
                    /rpt_family "AluJb"

Alignment Scores:
Pred. No.:                 29.8          Length:         147760
Score:                     51.00         Matches:        9
Percent Similarity:        100.00%       Conservative:   0
Best Local Similarity:     100.00%       Mismatches:     0
Query Match:               100.00%       Indels:         0
DB:                        9             Gaps:           0

SWOPE464-2 (1-9) x AC011846 (1-147760)

Qy          1 LeuAspGlnAsnValHisPheAlaHis 9
                |||||||||||||||||||||||||||
Db      16145 CTGGATCAGAATGTGCATTTTGCACAT 16169


Search completed: February 13, 2003, 14:32:44
Job time : 1804.12 secs
```

```
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Margulis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,S., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Raley,F., Roy,A., Santos,R., Seaman,S.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.-J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-OCT-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 117760)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Brown,J., Boguslavskiy,L.,
            Burkhardt,R., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chiaramonte,B., Chapel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Dhir,S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Gnerre,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Harris,K., Horton,L., Hulme,W., Iliev,I.,
            Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Margulis,N.,
            Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,L., Naylor,J., Nguyen,C.,
            Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
              Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Pollara,V., Raymond,C., Rettig,R., Riedberg,M., Riley,R., Rise,C.,
            Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
            Schupbach,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
            Theodore,J., Topham,K., Travers,M., Trazile,N., Tsiaris,T.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.-J.,
            Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUN-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 117760)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Brown,J., Boguslavskiy,L.,
            Burkhardt,R., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chiaramonte,B., Chapel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Dhir,S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Gnerre,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Harris,K., Horton,L., Hulme,W., Iliev,I.,
            Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Margulis,N.,
            Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,L., Naylor,J., Nguyen,C.,
            Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Pollara,V., Raymond,C., Rettig,R., Riedberg,M., Riley,R., Rise,C.,
            Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
            Schupbach,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
            Theodore,J., Topham,K., Travers,M., Trazile,N., Tsiaris,T.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.-J.,
            Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUN-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jun 17, 2002 this sequence version replaced [illegible].
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
                                Genome Center
              Center: Whitehead Institute/ MIT Center for Genome Research
              Center code: WIBR
              Web site: http://www-seq.wi.mit.edu
              Contact: sequence_submissions@genome.wi.mit.edu
                                Project Information
              Center project name: L2618
              Center clone name: 416_A_8
            [illegible]

            all the [illegible] 47.8 kb [illegible] this clone are being submitted.
            The [illegible] GenBank accession number AC[illegible] [illegible]
            [illegible].
FEATURES             Location/Qualifiers
     source          1..117760
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15"
                     /clone="RP11-416A8"
                     /clone_lib="RPCI-11 Human Male BAC"
     repeat_region   [illegible]..[illegible]
                     /rpt_family="MIR"
     repeat_region   [illegible]..[illegible]
                     /rpt_family="L1MB6"
     repeat_region   [illegible]..[illegible]
                     /rpt_family="AluJo"
     repeat_region   [illegible]..[illegible]
                     /rpt_family="AluSq"
     repeat_region   [illegible]..[illegible]
                     /rpt_family="L2"
     repeat_region   [illegible]..[illegible]
                     /rpt_family="AluJo"
     repeat_region   [illegible]..[illegible]
                     /rpt_family="(TG)n"
     repeat_region   [illegible]..[illegible]
                     /rpt_family="(CA)n"
     repeat_region   [illegible]..[illegible]
                     /rpt_family="(TG)n"
     repeat_region   complement([illegible]..[illegible])
                     /rpt_family="MIRs"
     repeat_region   [illegible]..[illegible]
                     /rpt_family="MIRs"
     repeat_region   complement([illegible]..[illegible])
                     /rpt_family="L2"
     repeat_region   [illegible]..[illegible]
                     /rpt_family="MIR"
     repeat_region   complement([illegible]..[illegible])
                     /rpt_family="AluSx"
     repeat_region   complement([illegible]..[illegible])
                     /rpt_family="AluSx"
     repeat_region   complement([illegible]..[illegible])
                     /rpt_family="MIR"
     repeat_region   [illegible]..[illegible]
                     /rpt_family="MIR"
     repeat_region   complement([illegible]..[illegible])
                     /rpt_family="AluY"
     repeat_region   complement([illegible]..[illegible])
                     /rpt_family="AluSx"
     repeat_region   [illegible]..[illegible]
                     /rpt_family="L2"
     repeat_region   [illegible]..[illegible]
                     /rpt_family="MIR"
     repeat_region   complement([illegible]..[illegible])
                     /rpt_family="L2"
     repeat_region   [illegible]..[illegible]
                     /rpt_family="MIRs"
     repeat_region   complement([illegible]..[illegible])
                     /rpt_family="AluJo"
     repeat_region   [illegible]..[illegible]
                     /rpt_family="MIRs"
     repeat_region   complement([illegible]..[illegible])
                     /rpt_family="AluSx"
     repeat_region   [illegible]..[illegible]
                     /rpt_family="L2"
     repeat_region   [illegible]..[illegible]
                     /rpt_family="L2"
     repeat_region   [illegible]..[illegible]
                     /rpt_family="(CA)n"
     repeat_region   [illegible]..[illegible]
```

```
Number of N's in consensus: [illegible]

                    Phrap Value Ratio

FEATURES             Location/Qualifiers
     source          1..[illegible]
                     /organism "[illegible]"
                     /strain "[illegible]"
                     /note "[illegible]"
     STS             [illegible]
                     /standard_name "[illegible]"
     repeat_region   258..[illegible]
                     /rpt_family "[illegible]"
     STS             [illegible]
                     /standard_name "[illegible]"
     repeat_region   complement([illegible])
                     /rpt_family "B4A"
     repeat_region   [illegible]
                     /rpt_family "[illegible]"
     repeat_region   complement([illegible])
                     /rpt_family "[illegible]"
     repeat_region   [illegible]
                     /rpt_family "MIR"
     repeat_region   [illegible]
                     /rpt_family "[illegible]"
     repeat_region   complement([illegible])
                     /rpt_family "[illegible]"
     repeat_region   [illegible]
                     /rpt_family "[illegible]"
     repeat_region   complement([illegible])
                     /rpt_family "MIR"
     repeat_region   [illegible]
                     /rpt_family "[illegible]"
     repeat_region   [illegible]
                     /rpt_family "[illegible]"
     repeat_region   complement([illegible])
                     /rpt_family "[illegible]"
     repeat_region   complement([illegible])
                     /rpt_family "[illegible]"
     repeat_region   [illegible]
                     /rpt_family "[illegible]"
     repeat_region   complement([illegible])
                     /rpt_family "[illegible]"
     repeat_region   complement([illegible])
                     /rpt_family "[illegible]"
     repeat_region   complement([illegible])
                     /rpt_family "[illegible]"
     repeat_region   complement([illegible])
                     /rpt_family "[illegible]"
     repeat_region   [illegible]
                     /rpt_family "[illegible]"
     repeat_region   complement([illegible])
                     /rpt_family "[illegible]"
     repeat_region   complement([illegible])
                     /rpt_family "MIR"
```

[illegible]

[illegible]
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
[illegible]
Millennium Pharmaceuticals, Inc. (US)
[illegible]
/note="unnamed protein product"
/translation="MAAAMLIE[illegible]VE[illegible]KLEEEYVE[illegible]
[illegible]
[illegible]
ORIGIN
[illegible]
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[illegible]
High-throughput Mouse Sequencing
Unpublished
[illegible]

[illegible]

REFERENCE [illegible]
AUTHORS [illegible] Montgomery,K.T., Brown,W.A., [illegible] Kucherlapati,R.
TITLE Direct Submission
JOURNAL Submitted [illegible] MAY [illegible]) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St., Cambridge, MA [illegible], USA

REFERENCE [illegible]
AUTHORS [illegible] Montgomery,K.T., Brown,W.A., [illegible] Kucherlapati,R.
TITLE Direct Submission
JOURNAL Submitted [illegible] JUN [illegible]) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St., Cambridge, MA [illegible], USA

REFERENCE [illegible]
AUTHORS [illegible] Montgomery,K.T., Brown,W.A., [illegible] Kucherlapati,R.
TITLE Direct Submission
JOURNAL Submitted [illegible] JUN [illegible]) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St., Cambridge, MA [illegible], USA

COMMENT [illegible] this sequence version replaced [illegible]
Genome Center:
Center: Harvard Partners Genome Center
[illegible]
Web site: http://www.[illegible]
[illegible]

[illegible] This sequence represents the entire [illegible]

ANNOTATION OF FEATURES:
[illegible] identified using [illegible] (Genome Res. [illegible])
[illegible] are identified using RepeatMasker (A. Smit [illegible]) for Human and Mouse sequences.
[illegible] regions of sequence similarity are identified [illegible] (Nucl. Acids Res. 25:3389-3402) similarity (expect [illegible]) to [illegible] and cDNA [illegible] in Unigene. [illegible] demonstrate at least [illegible] spliced [illegible] are [illegible] as similar.

SEQUENCING [illegible] COVERAGE: Attempts are made to complete [illegible] for all regions. All sequence is [illegible] with a minimum of [illegible] reads with [illegible] If the sequence [illegible] for a region does not meet this standard [illegible] in the annotation below. [illegible] are verified by PCR product size verification [illegible] forward and reverse reads from [illegible]

QUALITY [illegible] BASES: This sequence meets [illegible] quality standards. [illegible] estimated average error rate is less than [illegible] quality parameters. Regions that [illegible] are annotated [illegible] low quality.

Summary Statistics
[illegible]: AL1
Sequence [illegible] p2MAS[illegible] AL[illegible]
[illegible]

```
REFERENCE   2  (bases 1 to [illegible])
  AUTHORS   Abbott,C.A., Yu,D., McCaughan,G.W. and Gorrell,M.[illegible]
  TITLE     Direct Submission
  JOURNAL   Submitted ([illegible]-JAN-[illegible]) A.W. M[illegible] [illegible]
            [illegible] Centenary Institute of [illegible] and [illegible]
            Locked Bag No.[illegible], Newtown, Sydney, NSW 2042, Australia
FEATURES             Location/Qualifiers
     source          1..[illegible]
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1[illegible]"
                     /map="1[illegible]q22"
     gene            1..[illegible]
                     /gene="[illegible]"
     CDS             [illegible]
                     /gene="DPP8"
                     /note="serine protease; member of prolyl oligopeptidase
                     family belonging to [illegible] alternatively [illegible]
                     [illegible] this mRNA [illegible] AF22[illegible], AF22[illegible] and
                     AF22[illegible]7"
                     /codon_start=1
                     /product="dipeptidyl peptidase 8"
                     /protein_id="AA[illegible]"
                     /db_xref="[illegible]"
                     /translation="MAAAMET[illegible]
                     [illegible]
                     GEHYELHL[illegible]SR[illegible]AA[illegible]KVI"
BASE COUNT    [illegible]
ORIGIN

Alignment Scores:
Pred. No.:               [illegible]        Length:        [illegible]
Score:                   [illegible]        Matches:       [illegible]
Percent Similarity:      [illegible]        Conservative:  [illegible]
Best Local Similarity:   [illegible]        Mismatches:    [illegible]
Query Match:             [illegible]        Indels:        [illegible]
DB:                      [illegible]        Gaps:          [illegible]

SW [illegible] (1-[illegible]) x AF[illegible] (1-[illegible])

Qy    [illegible]LeuAsp[illegible]AsnVal[illegible]Phe[illegible]
Db    [illegible]

RESULT [illegible]
[illegible]
LOCUS       [illegible]        [illegible] bp    mRNA    linear   [illegible]
DEFINITION  Homo sapiens, Similar to dipeptidylpeptidase [illegible]
            IMAGE:[illegible], mRNA, complete cds.
ACCESSION   [illegible]
VERSION     [illegible]  GI:[illegible]
KEYWORDS    MGC.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
            Mammalia; Eutheria; Primates; [illegible]
REFERENCE   1  (bases 1 to [illegible])
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAY-[illegible]) National Institutes of Health, [illegible]
```

[illegible]

```
Db:                         Gaps:

Qy        [illegible]AspGlnAsnValHisPheAlaHis[illegible]

RESULT [illegible]
DEFINITION  Sequence [illegible] from Patent W[illegible].
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Meyers,R.A. and Williams,L.M.
  JOURNAL   Patent: WO [illegible]
            Millennium Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                       Length:
Score:                           Matches:
Percent Similarity:   100.00%    Conservative:
Best Local Similarity:           Mismatches:
Query Match:                     Indels:
DB:                              Gaps:

Qy        [illegible]AspGlnAsnValHisPheAlaHis[illegible]

RESULT [illegible]
AX[illegible]
DEFINITION  Sequence [illegible] from Patent WO[illegible].
ACCESSION   AX[illegible]
VERSION     AX[illegible]
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   [illegible]
  TITLE     Novel nucleic acids and polypeptides
  JOURNAL   Patent: WO 0222660-A [illegible] 21-MAR-2002;
            HYSEQ, INC. (US)
FEATURES             Location/Qualifiers
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="unnamed protein product"
                     /codon_start=1
                     /db_xref="[illegible]"
                     /translation="MAAAML[illegible]"
```

BASE COUNT [illegible]
ORIGIN

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:

RESULT [illegible]
AK[illegible]
LOCUS   AK[illegible]
DEFINITION [illegible]
ACCESSION   AK[illegible]
VERSION   AK[illegible]
KEYWORDS   [illegible]
SOURCE   [illegible]
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; [illegible] Vertebrata; [illegible]
REFERENCE   [illegible]
  AUTHORS   [illegible]
  TITLE   NEDO human cDNA [illegible]
  JOURNAL   [illegible]
REFERENCE   [illegible]
  AUTHORS   [illegible]
  TITLE   [illegible]
  JOURNAL   [illegible]
COMMENT   NEDO human cDNA [illegible]
FEATURES   [illegible]

```
AX[illegible]
LOCUS       AX[illegible]                [illegible] bp    DNA     linear   PAT [illegible]
DEFINITION  Sequence [illegible] from Patent WO[illegible]
ACCESSION   AX[illegible]
VERSION     AX08[illegible]
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to [illegible])
  AUTHORS   Yao,H., Lal,P., [illegible], Bandman,O., [illegible], Arvizu,C.,
            Lu,D.A. and Yue,H.
  TITLE     Proteases and protease inhibitors
  JOURNAL   Patent: WO [illegible]-A [illegible] 15-FEB-[illegible]
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..[illegible]
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: [illegible]"
BASE COUNT   [illegible]
ORIGIN

Alignment Scores:
Pred. No.:              [illegible]        Length:          [illegible]
Score:                  [illegible]        Matches:         [illegible]
Percent Similarity:     [illegible]        Conservative:    [illegible]
Best Local Similarity:  100.0%             Mismatches:      [illegible]
Query Match:            [illegible]        Indels:          [illegible]
DB:                     [illegible]        Gaps:            [illegible]

SW[illegible] (1-[illegible]) x AX[illegible] (1-[illegible])

Qy        [illegible]LeuAspGluAsnValHisPheAlaHis [illegible]

Db      411 [illegible]GGATGAGAA[illegible]CATTT[illegible]A[illegible]AT [illegible]37

RESULT 2
AF[illegible]
LOCUS       AF22[illegible]                1197 bp    mRNA    linear   PRI [illegible]
DEFINITION  Homo sapiens dipeptidyl peptidase [illegible] mRNA, partial cds,
            alternatively spliced.
ACCESSION   AF22[illegible]
VERSION     AF22[illegible].1  GI:[illegible]
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1197)
  AUTHORS   Abbott,C.A., Yu,D.M., [illegible] and McCaughan,G.W.
            and Gorrell,M.D.
  TITLE     [illegible] expression and chromosomal [illegible]
            dipeptidyl peptidase (DPP) IV homolog, DPP8
  JOURNAL   Eur. J. Biochem. 267 (2[illegible]), [illegible]
  MEDLINE   20497[illegible]
   PUBMED   [illegible]
REFERENCE   2  (bases 1 to 1197)
  AUTHORS   Abbott,C.A., Yu,D., McCaughan,G.W. and Gorrell,M.D.
  TITLE     Direct Submission
  JOURNAL   Submitted ([illegible]-JAN-[illegible]) A.W. Morrow [illegible]
            Centre, Centenary Institute of Cell Biology and Cancer Medicine,
            [illegible] Newtown, Sydney, NSW [illegible], Australia
FEATURES             Location/Qualifiers
     source          1..1197
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1[illegible]"
                     /map="1[illegible]q22"
                     /clone="121"
                     /tissue_type="placenta"
                     /clone_lib="lambdaTriplEx[illegible]"
```

```
[illegible]
```

SUMMARIES

look

ALIGNMENTS

```
[illegible]
ACCESSION   [illegible]
VERSION     [illegible]
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; [illegible] Vertebrata; [illegible]
            Mammalia; Eutheria; Primates; [illegible] Hominidae; Homo.
REFERENCE   1  (bases 1 to [illegible])
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer [illegible]
            Tumor Gene Index
  JOURNAL   Unpublished [illegible]
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: [illegible]
            [illegible]
            IMAGE [illegible]
            [illegible]
FEATURES             Location/Qualifiers
     source          1..[illegible]
                     [illegible]
```

```
Db:                  [illegible]            Gaps: [illegible]

SW [illegible] (1 [illegible]) x AW[illegible] ([illegible])

Qy        [illegible]AspI[illegible]AsnV[illegible]HisHisA[illegible]
              [illegible]
Db     [illegible]

RESULT [illegible]
A[illegible]
LOCUS       AI24[illegible]7        577 bp    mRNA    linear   EST [illegible]
DEFINITION  qh[illegible] Soares NFL [illegible] Homo sapiens cDNA clone
            IMAGE:[illegible] [illegible] WP:K[illegible] [illegible]
            sequence.
ACCESSION   AI24[illegible]7
VERSION     AI24[illegible]7.1  GI:[illegible]
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 577)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available [illegible]
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1276   Std Error: [illegible]
            Seq primer: [illegible] from Gibco
            High quality sequence stop: 1[illegible].
FEATURES             Location/Qualifiers
     source          1..577
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:[illegible]"
                     /clone_lib="Soares NFL [illegible]"
                     /lab_host="DH10B"
                     /note="[illegible] Vector [illegible]
                     a modified polylinker; Site 1: NotI; Site 2: EcoRI;
                     Equal amounts of plasmid DNA from [illegible]
                     libraries ([illegible] NFL [illegible] and [illegible]
                     NCI_CGAP_[illegible]) were mixed [illegible]
                     [illegible] with HAP [illegible] DNA was [illegible]
                     tracer in a subtractive hybridization reaction. The [illegible]
                     was PCR-amplified cDNAs [illegible]
                     from the same [illegible] libraries [illegible]
                     [illegible]
                     [illegible]
                     Soares and M. Fatima Bonaldo [illegible]"
BASE COUNT  [illegible] a  [illegible] c  12[illegible] g  1[illegible] t  [illegible] others
ORIGIN

Alignment Scores:
Pred. No.:             [illegible]          Length:        [illegible]
Score:                 [illegible]          Matches:       [illegible]
Percent Similarity:    [illegible]          Conservative:  [illegible]
Best Local Similarity: [illegible]          Mismatches:    [illegible]
Query Match:           [illegible]          Indels:        [illegible]
DB:                    [illegible]          Gaps:          [illegible]

SW [illegible] (1 [illegible]) x AI24[illegible] ([illegible])

Qy        [illegible]AspI[illegible]AsnV[illegible]HisHisA[illegible]
              [illegible]
Db     [illegible]

RESULT [illegible]
BM[illegible]
LOCUS       BM[illegible]        [illegible] bp    mRNA    linear   EST [illegible]
DEFINITION  [illegible] cDNA [illegible]
```

```
            [illegible]
ACCESSION   BM[illegible]
VERSION     BM[illegible]
KEYWORDS    [illegible]
SOURCE      [illegible]
  ORGANISM  [illegible]
            [illegible]
REFERENCE   [illegible]
  AUTHORS   [illegible]
  TITLE     [illegible]
  JOURNAL   [illegible]
COMMENT     [illegible]
FEATURES             [illegible]
     source          [illegible]
                     /[illegible]
                     /note="[illegible]"
BASE COUNT  [illegible]
ORIGIN

Alignment Scores:
Pred. No.:             [illegible]
Score:                 [illegible]
Percent Similarity:    [illegible]
Best Local Similarity: [illegible]
Query Match:           [illegible]
DB:                    [illegible]

SW [illegible]

Qy        [illegible]

Db        [illegible]

RESULT [illegible]
BI4[illegible]
```

BASE COUNT
ORIGIN

Alignment
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:

Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:

DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE human.
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES

BASE COUNT
ORIGIN

Alignment
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:

Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:

DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE human.
ORGANISM
REFERENCE
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
JOURNAL
COMMENT
FEATURES Location/Qualifiers
Equal amounts of plasmid DNA from three normal
BASE COUNT
ORIGIN

Alignment
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:

Length:
Matches:
Conservative:
Mismatches:
Indels:

```
                    constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT   [illegible]
ORIGIN

Alignment Scores:
Pred. No.:                 3.86          Length:        [illegible]
Score:                     [illegible]   Matches:       [illegible]
Percent Similarity:        [illegible]   Conservative:  [illegible]
Best Local Similarity:     100.00%       Mismatches:    0
Query Match:               100.00%       Indels:        [illegible]
DB:                        [illegible]   Gaps:          [illegible]

SW-HE464-1 (1-9) x AA278625 (1-4[illegible])

Qy    [illegible]AspIleAsnValHisLeuAla[illegible]
Db    [illegible]GGATGAGAATGCCATTTGCACAT 35[illegible]

RESULT [illegible]
AA[illegible]
LOCUS       AA[illegible]       50[illegible] bp    mRNA    linear   EST 18-AUG-[illegible]
DEFINITION  [illegible] NCI_CGAP_[illegible] Homo sapiens cDNA clone IMAGE:[illegible]
            similar to SW:[illegible]_HUMAN P42658 DIPEPTIDYL PEPTIDASE IV LIKE
            PROTEIN ;, mRNA sequence.
ACCESSION   AA[illegible]
VERSION     AA[illegible]  GI:[illegible]
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50[illegible])
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (199[illegible])
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allen,
            Ph.D., Gerald Marti, M.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            [illegible]
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: [illegible]
FEATURES             Location/Qualifiers
     source          1..5[illegible]
                     /organism="Homo sapiens"
                     /db_xref="GDB:[illegible]"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:[illegible]"
                     /clone_lib="NCI_CGAP_GCB1"
                     /tissue_type="germinal center B cell[illegible]"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA
                     was prepared from human [illegible]
                     germinal center B cells [illegible]
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allen
                     (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
                     primed with a NotI [illegible] primer
                     [illegible]TTTTTTTTTTTTTTTTTTTT[illegible]
                     Double-stranded cDNA was ligated to EcoRI adaptors
                     (Pharmacia), digested with NotI and cloned into the NotI
                     and EcoRI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
```

```
BASE COUNT   [illegible]
ORIGIN

Alignment Scores:
Pred. No.:                 [illegible]
Score:                     [illegible]
Percent Similarity:        [illegible]
Best Local Similarity:     [illegible]
Query Match:               [illegible]
DB:                        [illegible]

SW-HE464-1 [illegible]

Qy    [illegible]
Db    [illegible]

RESULT [illegible]
[illegible]
LOCUS       [illegible]   mRNA   [illegible]
DEFINITION  [illegible]
ACCESSION   [illegible]
VERSION     [illegible]
KEYWORDS    [illegible]
SOURCE      [illegible]
  ORGANISM  [illegible]
REFERENCE   [illegible]
  AUTHORS   [illegible]
  TITLE     [illegible]
  JOURNAL   Unpublished [illegible]
COMMENT     [illegible]
FEATURES    [illegible]
     source [illegible]
```

ORIGIN

similar to TR:O9*** DIPEPTIDYL PEPTIDASE IV. ;, mRNA sequence.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

JOURNAL Unpublished (1997)

ORIGIN

similar to TR:O9*** DIPEPTIDYL PEPTIDASE IV. ;, mRNA sequence.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

JOURNAL Unpublished (1997)

provided by Dr. Louis M. Staudt (NCI), Dr. David Allen (NCI) and Dr. Gerald Marti (CBER).

```
            [illegible]
            ,B., Morris,M., Parsons,J., [illegible], Pitkin,L., Rohlfing,T.,
            [illegible]
            [illegible] R.K., Wohldmann,P., Waterston,R.H., Wilson,R.K. and Marra,M.
  TITLE     Generation and analysis of 28[illegible] human expressed sequence tags
  JOURNAL   Genome Res. [illegible]
  MEDLINE   [illegible]
  COMMENT   Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box [illegible], St. Louis, MO [illegible]
            Tel: [illegible]
            Fax: [illegible]
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: [illegible]
            High quality sequence stop: [illegible]
FEATURES             Location/Qualifiers
     source          [illegible]
                     /organism="Homo sapiens"
                     /db_xref="[illegible]"
                     /db_xref="[illegible]"
                     /clone="IMAGE:[illegible]"
                     /clone_lib="[illegible]"
                     /lab_host="SOLR cells (kanamycin resistant)"
                     /note="[illegible]
                     [illegible]
                     [illegible] Pancreatic [illegible]
                     [illegible]
                     [illegible] adaptor sequence[illegible]
                     [illegible]"
BASE COUNT  [illegible]
ORIGIN

Alignment Scores:
Pred. No.:             [illegible]     Length:       [illegible]
Score:                 [illegible]     Matches:      [illegible]
Percent Similarity:    [illegible]     Conservative: [illegible]
Best Local Similarity: [illegible]     Mismatches:   [illegible]
Query Match:           [illegible]     Indels:       [illegible]
DB:                    9               Gaps:         [illegible]

[illegible] x AA[illegible]

[illegible] AspGlnAsnValHisProAlaHis[illegible]
           [illegible]
[illegible] AA[illegible]

RESULT 2
AA[illegible]
LOCUS       AA[illegible]     45[illegible] bp    mRNA    [illegible]   EST   [illegible]
DEFINITION  [illegible] mRNA
            sequence.
ACCESSION   AA[illegible]
VERSION     AA[illegible]
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; [illegible]
REFERENCE   1  (bases 1 to 45[illegible])
  AUTHORS   Adams,M.D., Kerlavage,A.R., [illegible]
            [illegible], Lee,N.H., Kirkness,E.F., [illegible]
            [illegible], Sutton,G., Blake,J.A., Brandon,R.[illegible]
            [illegible]
            [illegible], FitzHugh,W.M., [illegible], Nandhan,N.S., [illegible]
            [illegible]
            Kelley,[illegible], Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nelson,D.T., [illegible]
            Phillips,C.A., Ryder,S.E., Scott,J.L., [illegible]
            [illegible], Spriggs,T.A., [illegible]
            Bednarik,[illegible], Sepeda,M.A., [illegible]

[illegible]
```

SUMMARIES

ALIGNMENTS

RESULT
DEFINITION
ACCESSION
VERSION
KEYWORDS
ORGANISM
REFERENCE
AUTHORS

[illegible]

XX
[illegible] Human dipeptidyl peptidase (DPP) cDNA. [illegible]
XX
KW [illegible]
KW [illegible]
KW [illegible]
KW [illegible]
XX
[illegible]
XX
[illegible]
XX
[illegible] MAY [illegible]
XX
[illegible]
XX
[illegible]
XX
[illegible]
XX
[illegible]
XX
[illegible]
XX
PT New dipeptidyl peptidase (DPP) peptides, useful for screening [illegible]
PT [illegible]
PT [illegible]
PT [illegible]
XX
PS Example [illegible] English.
XX
CC [illegible] peptidase [illegible]
CC [illegible] such proteins. The DPP peptides are [illegible]
CC screening inhibitors of DPP catalytic activity. The inhibitors are useful
CC for treating [illegible] type II diabetes, cirrhosis, autoimmunity [illegible]
CC [illegible] HIV (human immunodeficiency virus) infection. The present
CC [illegible] cDNA.
XX
SQ Sequence [illegible]

Alignment Scores: [illegible]

[illegible]

```
XX
PN   W[illegible]A1.
XX
PD   [illegible] MAR [illegible]
XX
PF   [illegible] 2[illegible] W[illegible]
XX
PR   [illegible] 2000; [illegible]
XX
PA   [illegible] HYSEQ INC.
XX
PI   [illegible] A[illegible] V[illegible] [illegible]
PI   Xue A[illegible] Yang Y[illegible] Wehrman [illegible] [illegible]
XX
DR   WPI; [illegible]
DR   [illegible] ABK[illegible]
XX
PT   An isolated polynucleotide for treating diseases associated with the
PT   encoded polypeptide such as cancer and multiple sclerosis
XX
PS   [illegible] SEQ ID NO [illegible] English.
XX
CC   The present invention provides the protein and [illegible] sequences [illegible]
CC   [illegible] human proteins. These were isolated [illegible] expressed sequence [illegible]
CC   (ESTs). They can be used to stimulate cell growth [illegible]
CC   haematopoiesis e.g. to treat aplastic anaemia, to help tissue [illegible]
CC   [illegible] treatment, [illegible] regulate the immune system [illegible] to treat
CC   multiple sclerosis, [illegible] activity [illegible] to treat
CC   [illegible], to regulate haemostasis [illegible] to treat
CC   stroke and cancer, to screen for drugs, to treat inflammatory [illegible]
CC   e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC   Parkinson's disease. The present sequence is [illegible] sequence [illegible]
CC   invention.
XX
SQ   Sequence 24[illegible] BP; [illegible] A; [illegible] C; [illegible] G; [illegible] T; [illegible] other;

Alignment Scores:
Pred. No.:              [illegible]           Length:        24[illegible]
Score:                  [illegible]           Matches:       [illegible]
Percent Similarity:     [illegible]           Conservative:  [illegible]
Best Local Similarity:  [illegible]           Mismatches:    [illegible]
Query Match:            [illegible]           Indels:        [illegible]
DB:                     24                    Gaps:          [illegible]

SW [illegible] x ABN[illegible] (1-24[illegible])

Qy        [illegible]AspIleAsnValHis[illegible]AlaIle[illegible]
             [illegible]
Db   [illegible] [illegible]CAT[illegible]A[illegible]AAT[illegible]C[illegible]A[illegible]A[illegible] 24[illegible]5

RESULT [illegible]
ABK[illegible]
ID   ABK[illegible] standard; DNA; [illegible] BP.
XX
AC   ABK[illegible];
XX
DT   [illegible] Apr 2[illegible] (first entry)
XX
DE   Human protease PRISM cDNA sequence.
XX
KW   Human; protease; PRISM; gastrointestinal [illegible] disease; cancer;
KW   [illegible]; atherosclerosis; [illegible]
KW   inflammatory disorder; [illegible]
KW   [illegible] proliferative disorder; [illegible]
KW   Duchenne muscular dystrophy; epithelial [illegible]
KW   reproductive disorder; endometriosis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             [illegible]
FT                   /*tag= a
FT                   /product= "Human protease [illegible]"
```

```
XX
PN   W[illegible]A1.
XX
PD   [illegible]
XX
PF   [illegible] W[illegible]
XX
PR   [illegible]
PR   [illegible]
PR   [illegible]
PR   [illegible]
PR   [illegible]
XX
PA   (INCY ) INCYTE GENOMICS INC.
XX
PI   [illegible]
PI   [illegible]
PI   [illegible]
PI   [illegible]
PI   [illegible]
XX
DR   [illegible]
DR   [illegible]
XX
PT   [illegible]
PT   [illegible]
PT   [illegible]
PT   [illegible]
XX
PS   [illegible]
XX
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   myocardial infarction [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
XX
SQ   [illegible]

Alignment Scores:
Pred. No.:              [illegible]
Score:                  [illegible]
Percent Similarity:     [illegible]
Best Local Similarity:  [illegible]
Query Match:            [illegible]
DB:                     [illegible]

SW [illegible]

Qy   [illegible]

Db   [illegible]

RESULT [illegible]
AAS[illegible]
ID   AAS[illegible]
XX
AC   AAS[illegible];
XX
DT   [illegible]
XX
DE   [illegible]
XX
KW   [illegible]
KW   [illegible]
```

XX
KW [illegible] anti-inflammatory; [illegible]
KW [illegible]
KW [illegible]
KW [illegible]
XX
[illegible]
XX
PA (HYSE-) HYSEQ INC.
XX
PS Claim 1; SEQ ID NO 1695; [illegible]pp; English.
XX
CC The present invention provides the protein and [illegible] of [illegible] human proteins. These were isolated [illegible] (ESTs). They can be used to stimulate cell growth, to regulate haematopoiesis e.g. to treat aplastic anaemia, to help tissue repair with e.g. in burn treatment, to regulate the immune system e.g. to treat [illegible], to regulate [illegible], to treat infertility, to regulate haemostasis or [illegible], to treat stroke and cancer, to screen for drugs, to treat inflammatory conditions e.g. rheumatoid arthritis, and to treat nervous system disorders e.g. Parkinson's disease. The present sequence is [illegible]
XX
SQ Sequence [illegible]

Alignment Scores:
[illegible]

[illegible]

[illegible]

ABN[illegible]
[illegible]
XX
AC ABN[illegible]
XX
DT [illegible] (first entry)
XX
[illegible]
XX
KW [illegible] dipeptidyl peptidase IV [illegible]
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW [illegible]
KW [illegible]
KW [illegible]
KW [illegible]

KW [illegible]
XX
OS Homo sapiens.
XX
PN WO[illegible]
XX
PD [illegible]
XX
[illegible]
XX
[illegible]
XX
PA (FERR-) FERRING BV.
XX
[illegible]
XX
DR WPI; [illegible]
DR P-PSDB; ABN[illegible]
XX
PT New [illegible]
PT [illegible]
PT [illegible] metabolic disorders [illegible]
PT
XX
PS [illegible] English.
XX
CC The present invention relates to the isolation of novel human serine proteases referred to as dipeptidyl peptidase IV (DPPIV) related proteins (DPRs). The dipeptidyl peptidase IV related proteins (DPRs) and [illegible] them are useful for treating infections such as fungal, [illegible], protozoan and viral infections, particularly infections caused by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes, precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's disease, acute heart failure, hypotension, hypertension, [illegible] retention, osteoporosis, angina pectoris, stroke, [illegible], allergies, cancers, migraine, vomiting, psychotic and neurological disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias. These may also be used in discovering therapeutic agents for the treatment of reproductive, inflammatory and metabolic [illegible] ABK[illegible] encode human DPR [illegible]
XX
SQ Sequence [illegible]

Alignment Scores:
[illegible]

[illegible]

[illegible]

[illegible]

RESULT [illegible]
ABN[illegible]
ID ABN[illegible]
XX
AC ABN[illegible]
XX
DT 28-JUN-[illegible]
XX
DE [illegible]
XX
KW Human; [illegible] anti-inflammatory; immunomodulator;
KW antidiabetic; [illegible]
KW [illegible]
KW [illegible]
XX
OS [illegible]

```
PI   [illegible]
PI   [illegible] AE, [illegible]
PI   Reddy [illegible]
XX
DR   WPI; [illegible]
DR   [illegible] AAB[illegible]
XX
PT   Novel human proteases and polynucleotides [illegible]
PT   useful for treating, diagnosing or preventing cell proliferative,
PT   [illegible] autoimmune/inflammatory, [illegible]
PT   disorders -
XX
PS   [illegible]
XX
CC   The invention relates to human proteases (PRTS [illegible]) and [illegible]
CC   cDNA molecules. Human PRTS and its [illegible]
CC   the diagnosis, treatment and prevention of disorders associated with
CC   increased or decreased expression of PRTS. Examples of such [illegible]
CC   include, cell proliferative disorders (arteriosclerosis, atherosclerosis,
CC   hepatitis, psoriasis and cancers); autoimmune/inflammatory disorders
CC   (AIDS, Addison's disease, [illegible] asthma, [illegible]
CC   diabetes mellitus, [illegible], multiple sclerosis, [illegible]
CC   trauma, Grave's disease, rheumatoid arthritis, [illegible]
CC   viral, bacterial, fungal, parasitic, protozoal and helminthic
CC   infections); cardiovascular disorders ([illegible]
CC   heart disease and hypertension); neurological disorders (epilepsy,
CC   Alzheimer's disease, Pick's disease, Huntington's disease, dementia,
CC   Parkinson's disease, stroke, mental disorders (including [illegible] anxiety
CC   and [illegible] disorder and prion diseases); gastrointestinal
CC   disorders (Crohn's disease, anorexia, [illegible], diarrhoea and [illegible]);
CC   epithelial disorders (contact dermatitis, eczema, acne, [illegible],
CC   alopecia, scabies, insect bites and urticaria); reproductive disorders
CC   (infertility, disruption of oestrous and [illegible] and
CC   [illegible] and developmental disorders ([illegible]
CC   Cushing's syndrome, seizure disorders, [illegible] and [illegible]).
CC   PRTS DNA is also in useful is gene therapy. PRTS and its immunogenic
CC   fragments are useful for screening libraries of compounds in screening
CC   [illegible] assays. The present sequence is human protease PRTS [illegible] cDNA.
XX
SQ   Sequence 2510 BP; 777 A; 104 C; 527 G; [illegible] other;

Alignment Scores:
Pred. No.:                [illegible]     Length:          [illegible]
Score:                    [illegible]     Matches:         [illegible]
Percent Similarity:       [illegible]     Conservative:    [illegible]
Best Local Similarity:    [illegible]     Mismatches:      [illegible]
Query Match:              [illegible]     Indels:          [illegible]
DB:                       24              Gaps:            [illegible]

SW [illegible] x AAB[illegible] ([illegible])

Qy        [illegible]AspLeuAsnValHisProAlaHis[illegible]

Db   [illegible]

RESULT [illegible]
AAB[illegible]
ID   AAB[illegible] standard; cDNA; [illegible]
XX
AC   AAB[illegible];
XX
DT   [illegible]
XX
DE   [illegible] sequence of [illegible] human prolyl [illegible]peptidase.
XX
KW   [illegible] prolyl [illegible] peptidase; [illegible]
KW   [illegible] cardiovascular disease; [illegible]
KW   neuronal disorder; vascular disorder; [illegible] disorder; [illegible]
KW   antidiabetic; antiarthritic; antiasthmatic; [illegible]
KW   diabetes mellitus; arthritis; multiple [illegible] asthma;
KW   Grave's disease; [illegible] disorder; [illegible]
XX
OS   Homo sapiens.
```

```
XX
PN   W[illegible]
XX
PD   [illegible]
XX
PF   [illegible]
XX
PR   [illegible]
XX
PA   [illegible]
XX
PI   [illegible]
XX
DR   [illegible]
DR   [illegible]
XX
PT   Novel [illegible]
PT   [illegible]
PT   [illegible]
XX
PS   [illegible]
XX
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   antiparkinsonian, [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   differentiation [illegible]
CC   [illegible]
CC   [illegible]
CC   proliferative and differentiation [illegible]
CC   disorders [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   AAB[illegible].
XX
SQ   [illegible]

Alignment Scores:
Pred. No.:                [illegible]
Score:                    [illegible]
Percent Similarity:       [illegible]
Best Local Similarity:    [illegible]
Query Match:              [illegible]
DB:                       [illegible]

SW [illegible]

Qy   [illegible]

Db   [illegible]

RESULT [illegible]
ABN[illegible]
ID   ABN[illegible]
XX
AC   ABN[illegible]
XX
DT   [illegible]
XX
DE   N[illegible]
```

XX

Human cDNA sequence ...

XX

...

Human protease ... cDNA.

XX

KW ... atherosclerosis; rheumatoid arthritis ...

KW ... Alzheimer's disease; mental disorder ...

KW ... hypertension ... Parkinson's disease ...

KW ... multiple sclerosis ...

KW ... myocardial infarction ...

KW ... infertility ...

XX

... "Human protease ... protein"

XX

XX

PA (INST...

```
[illegible]

ID   AA[illegible] standard; cDNA; [illegible]
XX
AC   AA[illegible]
XX
[illegible]
XX
[illegible]
XX
KW   Human; dipeptidyl aminopeptidase [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
XX
[illegible]
XX
[illegible]
XX
[illegible]
XX
FA   (UNSY ) UNIV SYDNEY.
XX
[illegible]
XX
[illegible]
DR   [illegible] AAB47[illegible]
XX
PT   New human dipeptidyl amino peptidase (DPP[illegible]) [illegible]
PT   substrates, identifying inhibitors of [illegible]
PT   [illegible]
XX
[illegible]
XX
CC   The sequences given in AA[illegible] encode fragments of human
CC   dipeptidyl amino peptidase (DPP[illegible]). [illegible]
CC   H-Ala-Pro-pNA, H-Gly-Pro-pNA and H-Arg-Pro-pNA [illegible]
CC   [illegible]
```

[illegible]

XX

[illegible] Human [illegible]

XX

[illegible]

XX

[illegible] cardiovascular, renal and [illegible]

XX

[illegible]

XX

[illegible] novel genes (ABR [illegible]) and [illegible] gene therapy. [illegible] treatment and prevention of: (a) cancer [illegible] intestinal tract, [illegible], lung, [illegible] (b) [illegible] Addison's disease, allergies, [illegible] diabetes [illegible] rheumatoid arthritis and [illegible] (d) wound healing; (e) neurological diseases [illegible] cerebral [illegible] and [illegible].

Note: The [illegible] this patent did not form part of the printed [illegible] from WIPO [illegible] int/pub/published pct sequences.

XX

[illegible]

```
XX
PR   [illegible] JAN [illegible]
PR   [illegible] MAY [illegible]
XX
PA   [illegible]
XX
PI   [illegible] Nishikawa [illegible] Hayashi [illegible]
PI   Ishii S, Sugiyama [illegible], Wakamatsu A, Nagai K [illegible]
XX
DR   WPI; [illegible]
XX
PT   Primer sets for synthesizing polynucleotides, particularly the
PT   full length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full length cDNAs
XX
PS   [illegible] SEQ ID [illegible]; English.
XX
CC   The present invention describes primer sets for synthesising
CC   full length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides
CC   particularly full length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full length cDNAs. The primers allow [illegible] of the full length
CC   cDNAs easily without any specialised methods. AAH[illegible] to AAH[illegible] and
CC   AAH[illegible] to AAH18742 represent human cDNA sequences; AAB[illegible] to
CC   AAB[illegible] represent human amino acid sequences; and AAH1[illegible] to AAH[illegible]
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence [illegible]

Alignment Scores:
Pred. No.:              [illegible]      Length:         [illegible]
Score:                  [illegible]      Matches:        [illegible]
Percent Similarity:     [illegible]      Conservative:   [illegible]
Best Local Similarity:  [illegible]      Mismatches:     [illegible]
Query Match:            [illegible]      Indels:         [illegible]
DB:                     22               Gaps:           [illegible]

[illegible] AAH[illegible]

RESULT [illegible]
AAB[illegible]
ID   AAB[illegible] standard; [illegible]
XX
AC   AAB[illegible]
XX
DT   [illegible] (first entry)
XX
DE   Human [illegible] protease [illegible]
XX
KW   Human; [illegible] therapy;
KW   [illegible]; protease mediated disorder; proliferative disorder;

KW   [illegible]
XX
DT   [illegible] MAR [illegible]
XX
PR   [illegible] MAR [illegible]
XX
PA   [illegible]
XX
[illegible]
XX
Alignment Scores:
Pred. No.:              [illegible]
Score:                  [illegible]
Percent Similarity:     [illegible]
Best Local Similarity:  [illegible]
Query Match:            [illegible]
DB:                     [illegible]

RESULT [illegible]
AAB[illegible]
ID   AAB[illegible]
```

GENERAL INFORMATION:
APPLICANT: [illegible]
TITLE OF INVENTION: [illegible]
NUMBER OF SEQUENCES: [illegible]
CORRESPONDENCE ADDRESS:
ADDRESSEE: [illegible]
STATE: New York
COUNTRY: United States of America
COMPUTER READABLE FORM:
MEDIUM TYPE: [illegible]
COMPUTER: IBM PC compatible
OPERATING SYSTEM: [illegible]
SOFTWARE: [illegible]
CURRENT APPLICATION DATA:
APPLICATION NUMBER: [illegible]
FILING DATE: [illegible] MAR [illegible]
CLASSIFICATION: [illegible]
PRIOR APPLICATION DATA:
APPLICATION NUMBER: [illegible]
FILING DATE: [illegible] December [illegible]
APPLICATION NUMBER: [illegible]
FILING DATE: [illegible] December [illegible]
ATTORNEY/AGENT INFORMATION:
NAME: [illegible]
REGISTRATION NUMBER: [illegible]
REFERENCE/DOCKET NUMBER: [illegible]
TELECOMMUNICATION INFORMATION:
TELEPHONE: [illegible]
TELEFAX: [illegible]
INFORMATION FOR SEQ ID NO: [illegible]
SEQUENCE CHARACTERISTICS:
LENGTH: [illegible] base pairs
TYPE: nucleic acid
STRANDEDNESS: [illegible]
TOPOLOGY: [illegible]
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
[illegible]

[illegible]

GENERAL INFORMATION:
APPLICANT: [illegible] ATHANASIOS
APPLICANT: [illegible] TAKAHIRO
TITLE OF INVENTION: [illegible]
TITLE OF INVENTION: [illegible]
NUMBER OF SEQUENCES: [illegible]
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
[illegible]
COUNTRY: USA

[illegible]
COMPUTER: IBM PC compatible
OPERATING SYSTEM: [illegible]
SOFTWARE: [illegible] Version [illegible]
CURRENT APPLICATION DATA:
APPLICATION NUMBER: [illegible]
FILING DATE:
CLASSIFICATION: [illegible]
PRIOR APPLICATION DATA:
APPLICATION NUMBER: [illegible]
FILING DATE: [illegible]
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: [illegible]
REFERENCE/DOCKET NUMBER: [illegible]
TELECOMMUNICATION INFORMATION:
TELEPHONE: [illegible]
TELEFAX: [illegible]
TELEX: [illegible]
INFORMATION FOR SEQ ID NO: [illegible]
SEQUENCE CHARACTERISTICS:
LENGTH: [illegible]
TYPE: [illegible]
STRANDEDNESS: [illegible]
TOPOLOGY: [illegible]
FEATURE:
NAME/KEY: [illegible]
LOCATION: [illegible]
[illegible]

Alignment Scores:
Pred. No.: [illegible] Length: [illegible]
Score: [illegible] Matches: [illegible]
Percent Similarity: [illegible]% Conservative: [illegible]
Best Local Similarity: [illegible]% Mismatches: [illegible]
Query Match: [illegible] Indels: [illegible]
DB: [illegible] Gaps: [illegible]

SWOPE464-1 [illegible]

[illegible]

[illegible]

Search completed: [illegible]
[illegible]

```
;       NAME/KEY: [illegible]
;       LOCATION: [illegible]
US-[illegible]

Alignment Scores:
Pred. No.:                 [illegible]      Length:          [illegible]
Score:                     43.00            Matches:         [illegible]
Percent Similarity:        [illegible]      Conservative:    [illegible]
Best Local Similarity:     [illegible]      Mismatches:      [illegible]
Query Match:               72.88%           Indels:          [illegible]
DB:                        [illegible]      Gaps:            [illegible]

SW[illegible] x US-[illegible]

Qy        [illegible]
Db        [illegible]

RESULT 12
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;     APPLICANT:  Vladimir Shamanin
;     APPLICANT:  Ethel Michele De Villiers
;     APPLICANT:  Zur Hausen
;     APPLICANT:  Irene Leigh
;     APPLICANT:  Harald Zur Hausen
;     TITLE OF INVENTION:  DNA CODING FOR A PEPTIDE OF A PAPILLOMA
;     TITLE OF INVENTION:  VIRUS MAIN CAPSID PROTEIN AND USE THEREOF
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Pennie & Edmonds LLP
;       STREET:  [illegible] Avenue of the Americas
;       CITY:  New York
;       STATE:  New York
;       COUNTRY:  USA
;       ZIP:  [illegible]
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy Disk
;       COMPUTER:  IBM PC Compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #[illegible] Version #[illegible]
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  [illegible]
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  [illegible]
;       FILING DATE:  2[illegible]-[illegible]AN-[illegible]
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Adams, Samuel B.
;       REGISTRATION NUMBER:  [illegible]
;       REFERENCE/DOCKET NUMBER:  [illegible]484-[illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  ([illegible]) 4[illegible]-[illegible]
;       TELEFAX:  ([illegible])4[illegible] 5556
;       TELEX:  66[illegible]14 PENNIE
;  INFORMATION FOR SEQ ID NO: [illegible]:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  [illegible] base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  [illegible]
US-[illegible]

Alignment Scores:
Pred. No.:                 [illegible]      Length:          [illegible]
Score:                     4[illegible]     Matches:         [illegible]

[illegible]
Query Match:               [illegible]
DB:                        [illegible]

SW[illegible]

Qy        [illegible]
Db        [illegible]

RESULT [illegible]
US-[illegible]
; Sequence [illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;     APPLICANT:  [illegible]
;     TITLE OF INVENTION:  [illegible]
;     TITLE OF INVENTION:  [illegible]
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  [illegible]
;       STREET:  [illegible]
;       CITY:  [illegible]
;       STATE:  New York
;       COUNTRY:  [illegible]
;       ZIP:  [illegible]
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  [illegible]
;       COMPUTER:  [illegible]
;       OPERATING SYSTEM:  [illegible]
;       SOFTWARE:  [illegible]
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  [illegible]
;       FILING DATE:  [illegible]
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  [illegible]
;       FILING DATE:  [illegible]
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  [illegible]
;       REGISTRATION NUMBER:  [illegible]
;       REFERENCE/DOCKET NUMBER:  [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  [illegible]
;       TELEFAX:  [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  [illegible]
;       TYPE:  [illegible]
;       STRANDEDNESS:  [illegible]
;       TOPOLOGY:  [illegible]
;     MOLECULE TYPE:  [illegible]
;     HYPOTHETICAL:  [illegible]
US-[illegible]

Alignment Scores:
Pred. No.:                 [illegible]
Score:                     [illegible]
Percent Similarity:        [illegible]
Best Local Similarity:     [illegible]
Query Match:               [illegible]
DB:                        [illegible]

SW[illegible]

Qy        [illegible]
Db        [illegible]

RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application [illegible]
; Patent No. [illegible]
```

```
[illegible]
COMPUTER READABLE FORM:
MEDIUM TYPE: [illegible]
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSEQ for Windows Version [illegible]
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/[illegible]
FILING DATE: [illegible]
CLASSIFICATION: [illegible]
PRIOR APPLICATION DATA:
APPLICATION NUMBER: [illegible]
FILING DATE: [illegible]
PRIOR APPLICATION DATA:
APPLICATION NUMBER: [illegible]
FILING DATE: [illegible] JAN [illegible]
PRIOR APPLICATION DATA:
APPLICATION NUMBER: [illegible]
FILING DATE: [illegible] APR [illegible]
PRIOR APPLICATION DATA:
APPLICATION NUMBER: [illegible]
FILING DATE: [illegible]
ATTORNEY/AGENT INFORMATION:
NAME: [illegible]
REGISTRATION NUMBER: [illegible]
REFERENCE/DOCKET NUMBER: [illegible]
TELECOMMUNICATION INFORMATION:
TELEPHONE: [illegible]
TELEFAX: [illegible]
TELEX: [illegible]
INFORMATION FOR SEQ ID NO: [illegible]
SEQUENCE CHARACTERISTICS:
LENGTH: [illegible] base pairs
TYPE: nucleic acid
STRANDEDNESS: [illegible]
TOPOLOGY: [illegible]
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: N
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: [illegible] MONAS [illegible] VALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: [illegible]
[illegible]

Alignment Scores:
[illegible]

[illegible]

RESULT [illegible]
[illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;   APPLICANT: [illegible]
;   APPLICANT: [illegible]
;   APPLICANT: [illegible]
;   APPLICANT: [illegible]
;   APPLICANT: [illegible]
;   TITLE OF INVENTION: DNA [illegible] FOR A PEPTIDE OF A [illegible]
;   TITLE OF INVENTION: VIRUS MAIN CAPSIDE PROTEIN AND USE THEREOF
;   NUMBER OF SEQUENCES: [illegible]
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: [illegible] Edmonds LLP
;     STREET: [illegible] Avenue of the Americas
;     CITY: New York
;     STATE: New York
;     COUNTRY: USA
;     ZIP: [illegible]
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy Disk
;     COMPUTER: IBM [illegible]
;     OPERATING SYSTEM: [illegible]
;     SOFTWARE: [illegible] Version [illegible]
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/[illegible]
;     FILING DATE: [illegible] JAN [illegible]
;     CLASSIFICATION: [illegible]
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: [illegible]
;     FILING DATE: [illegible]
;   ATTORNEY/AGENT INFORMATION:
;     NAME: [illegible]
;     REGISTRATION NUMBER: [illegible]
;     REFERENCE/DOCKET NUMBER: [illegible]
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: [illegible]
;     TELEFAX: [illegible]
;     TELEX: [illegible]
; INFORMATION FOR SEQ ID NO: [illegible]
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: [illegible]
;     TYPE: [illegible]
;     STRANDEDNESS: [illegible]
;     TOPOLOGY: [illegible]
;   MOLECULE TYPE: DNA
;   FEATURE:
```

```
Score:                     48.0         Matches:        [illegible]
Percent Similarity:        100.0%       Conservative:   [illegible]
Best Local Similarity:     87.5%        Mismatches:     [illegible]
Query Match:               81.36%       Indels:         0
DB:                        1            Gaps:           0

SW [illegible]

Qy    [illegible] TyrTrpSerTyrGlyGlyTyrLeu 9
                  [illegible]
Db    [illegible] TGTTGGAGTTATGGAGGATATCTT [illegible]

RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;    APPLICANT: Zimmermann, Rainer; Park, [illegible]
;    APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
;    TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION
;    TITLE OF INVENTION: PROTEIN ALPHA, AND USES THEREOF
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Felfe & Lynch
;      STREET: [illegible] Third Avenue
;      CITY: New York City
;      STATE: New York
;      COUNTRY: USA
;      ZIP: [illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Diskette, [illegible] inch, 2[illegible] Mb storage
;      COMPUTER: IBM PS/2
;      OPERATING SYSTEM: PC-DOS
;      SOFTWARE: Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/[illegible]
;      FILING DATE: [illegible] 199[illegible]
;      CLASSIFICATION: [illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 08/[illegible]
;      FILING DATE: [illegible]-MARCH-[illegible]
;      APPLICATION NUMBER: 08/[illegible]
;      FILING DATE: [illegible]-APRIL-[illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Hanson, Norman [illegible]
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: LUD [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (212) [illegible]
;      TELEFAX: (212) [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: [illegible] base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: double
;      TOPOLOGY: linear
US-[illegible]

Alignment Scores:
Pred. No.:                 [illegible]  Length:         [illegible]
Score:                     48.0         Matches:        [illegible]
Percent Similarity:        100.00%      Conservative:   [illegible]
Best Local Similarity:     87.5%        Mismatches:     [illegible]
Query Match:               81.36%       Indels:         [illegible]
DB:                        2            Gaps:           [illegible]

SW [illegible]

Qy    [illegible] TyrTrpSerTyrGlyGlyTyrLeu 9
                  [illegible]
Db    [illegible] TGTTGGAGTTATGGAGGATATCTT [illegible]

RESULT [illegible]
```

[illegible]

COMPUTER READABLE FORM:
MEDIUM TYPE:
OPERATING SYSTEM:
SOFTWARE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME:
APPLICATION NUMBER:
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE:
TELEFAX:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: Base pairs
TYPE:
STRANDEDNESS:
TOPOLOGY:

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match: 86.44%
DB: 4

Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:

RESULT
GENERAL INFORMATION:
APPLICANT:
APPLICANT:
TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES:
CORRESPONDENCE ADDRESS:
ADDRESSEE:
STREET: THIRD AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP:
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE
COMPUTER:
OPERATING SYSTEM:
SOFTWARE: WORDPERFECT
CURRENT APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE: 20 APRIL
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME:

REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE:
TELEFAX:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: Base pairs
TYPE:
STRANDEDNESS:
TOPOLOGY:

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:

Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:

RESULT
; Sequence 1, Application
; GENERAL INFORMATION:
; APPLICANT:
; APPLICANT:
; TITLE OF INVENTION:
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES:
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:
; STREET:
; CITY: New York
; STATE: New York
; COUNTRY:
; ZIP:
; COMPUTER READABLE FORM:
; MEDIUM TYPE:
; COMPUTER:
; OPERATING SYSTEM:
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:
; TELEFAX:
; INFORMATION FOR SEQ ID NO:
; SEQUENCE CHARACTERISTICS:
; LENGTH: Base pairs
; TYPE:
; STRANDEDNESS:
; TOPOLOGY:

Alignment Scores:
Pred. No.:
Length:

```
; [illegible] Application US/[illegible]
; [illegible]
; GENERAL INFORMATION:
;   APPLICANT: [illegible], Bruce [illegible]
;   TITLE OF INVENTION: [illegible]
;   NUMBER OF SEQUENCES: [illegible]
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: MORRISON & FOERSTER
;     STREET: [illegible] PAGE MILL ROAD
;     CITY: Palo Alto
;     STATE: CA
;     COUNTRY: USA
;     ZIP: [illegible]
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette
;     COMPUTER: IBM Compatible
;     OPERATING SYSTEM: Windows
;     SOFTWARE: FastSEQ for Windows Version [illegible]
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/[illegible]
;     FILING DATE: 23-DEC-199[illegible]
;     CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: [illegible]
;     FILING DATE: [illegible]-1997
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: [illegible]
;     FILING DATE: [illegible]-JAN-199[illegible]
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: [illegible]
;     FILING DATE: [illegible]-APR-1998
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: [illegible]
;     FILING DATE: [illegible]
;   ATTORNEY/AGENT INFORMATION:
;     NAME: [illegible]
;     REGISTRATION NUMBER: [illegible]
;     REFERENCE/DOCKET NUMBER: 27[illegible]-20[illegible]
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: [illegible]
;     TELEFAX: [illegible]
;     TELEX: [illegible]
; INFORMATION FOR SEQ ID NO: [illegible]
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: [illegible] base pairs
;     TYPE: [illegible]
;     STRANDEDNESS: [illegible]
;     TOPOLOGY: circular
;   MOLECULE TYPE: DNA (genomic)
;   HYPOTHETICAL: NO
;   ANTI-SENSE: UNKNOWN
;   ORIGINAL SOURCE:
;     ORGANISM: [illegible]
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: [illegible]
US-[illegible]

Alignment Scores:
Pred. No.: [illegible]   Length: [illegible]
Score: [illegible]   Matches: [illegible]
Percent Similarity: 100.00%   Conservative: [illegible]
Best Local Similarity: [illegible]   Mismatches: [illegible]
Query Match: [illegible]   Indels: [illegible]
DB: 4   Gaps: [illegible]
```

[illegible]

```
RESULT 4
; [illegible]
; GENERAL INFORMATION:
;   APPLICANT: [illegible]
;   TITLE OF INVENTION: [illegible]
;   NUMBER OF SEQUENCES: [illegible]
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: [illegible]
;     STREET: [illegible]
;     CITY: [illegible]
;     STATE: CA
;     COUNTRY: USA
```

Listing first 45 summaries

SUMMARIES

ALIGNMENTS

RESULT 1

; Sequence [illegible], Application US/[illegible]A
; Patent No. [illegible]
; GENERAL INFORMATION:
; APPLICANT: [illegible]
; TITLE OF INVENTION: [illegible] Human [illegible]
; FILE REFERENCE: [illegible]
; CURRENT APPLICATION NUMBER: US/[illegible]A
; CURRENT FILING DATE: [illegible]
; NUMBER OF SEQ ID NOS: [illegible]
; SOFTWARE: PatentIn Ver. [illegible]
; SEQ ID NO [illegible]
; LENGTH: [illegible]
; TYPE: DNA
; ORGANISM: [illegible]
; FEATURE:
; OTHER INFORMATION: [illegible]peptidases

Alignment Scores:

Pred. No.:	[illegible]	Length:	[illegible]
Score:	[illegible]	Matches:	[illegible]
Percent Similarity:	[illegible]	Conservative:	[illegible]
Best Local Similarity:	[illegible]	Mismatches:	[illegible]
Query Match:	[illegible]	Indels:	[illegible]
DB:	[illegible]	Gaps:	[illegible]

RESULT 2

```
            High quality sequence start: [illegible]
            High quality sequence stop: [illegible]
FEATURES             [illegible]
     source          [illegible]
                     /organism="Homo sapiens"
                     /db_xref="taxon:[illegible]"
                     /clone="IMAGE:[illegible]"
                     /clone_lib="NIH_MGC_[illegible]"
                     /tissue_type="epidermoid carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="[illegible] Vector: pBluescript [illegible]
                     Site_2: EcoRI; cDNA made by oligo-dT primed,
                     directionally cloned into EcoRI/XhoI sites using the
                     following [illegible] adapter: [illegible] library constructed
                     by [illegible] in the laboratory of Gerald M. Rubin
                     (University of California, Berkeley) using ZAP cDNA
                     synthesis kit (Stratagene) and SuperScript II RT (Life
                     Technologies). Note: this is a NIH_MGC Library."
BASE COUNT    [illegible] a   [illegible] c   [illegible] g   [illegible] t
ORIGIN

Alignment Scores:
Pred. No.:                 [illegible]          Length:          [illegible]
Score:                     [illegible]          Matches:         [illegible]
Percent Similarity:        [illegible]          Conservative:    [illegible]
Best Local Similarity:     [illegible]          Mismatches:      [illegible]
Query Match:               100.00%              Indels:          [illegible]
DB:                        [illegible]          Gaps:            [illegible]

SW [illegible] (1-[illegible]) x [illegible] (1-[illegible])

Qy       [illegible] HisTyrIleProGlyTyrGlyTyrTyrIle[illegible]
                          [illegible]
Db       [illegible] CATTCCTGGC[illegible]TACGGA[illegible]GATAC[illegible]

Search completed: February [illegible]
Job time : [illegible] secs
```

```
[illegible]
REFERENCE    [illegible]
  AUTHORS    NIH MGC [illegible]
  TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL    Unpublished [illegible]
COMMENT      Contact: Robert Strausberg, Ph.D.
             [illegible]
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             [illegible]
FEATURES             Location/Qualifiers
     source          [illegible]
                     /organism="Homo sapiens"
                     [illegible]
                     /clone="IMAGE:[illegible]"
                     /clone_lib="NIH_MGC_[illegible]"
                     [illegible]
                     /note="Organ: brain; Vector: pBluescriptR (modified
                     pBluescript KS); [illegible]
                     [illegible] library constructed
                     [illegible] using the
                     Cap-Trapper method (Carninci et al., in preparation). Library
                     constructed by M. Bonaldo [illegible]
                     [illegible] National
                     Institutes of Health; Note: this is a NIH_MGC Library."
BASE COUNT   [illegible]
ORIGIN

Alignment Scores:
[illegible]

[illegible]

REFERENCE    [illegible]
  AUTHORS    NIH MGC [illegible]
  TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL    Unpublished [illegible]
COMMENT      Contact: Robert Strausberg, Ph.D.
             [illegible]
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             [illegible]
FEATURES             Location/Qualifiers
     source          [illegible]
                     [illegible]
                     /note="Organ: [illegible] Vector: pCMV-SPORT6 [illegible]
                     [illegible] NIH_MGC Library."
BASE COUNT   [illegible]
ORIGIN

Alignment Scores:
[illegible]

SWOPE464-1 [illegible]

[illegible]

RESULT [illegible]
[illegible]
LOCUS        [illegible]
DEFINITION   [illegible] NIH_MGC [illegible] Homo sapiens cDNA clone IMAGE:[illegible]
             mRNA sequence.
ACCESSION    [illegible]
VERSION      [illegible]
KEYWORDS     EST.
SOURCE       [illegible]
  ORGANISM   [illegible]
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
REFERENCE    [illegible]
  AUTHORS    NIH MGC [illegible]
  TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL    Unpublished [illegible]
COMMENT      Contact: Robert Strausberg, Ph.D.
             [illegible]
             cDNA Library Preparation: [illegible]
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             [illegible]
```

```
                    /db_xref="[illegible]"
                    /clone="[illegible]"
                    /clone_lib="NIH MGC 7"
                    /tissue_type="[illegible]"
                    /[illegible]="[illegible]"
                    /lab_host="[illegible]"
                    /note="[illegible] Vector: [illegible] Site 1: XbaI [illegible]
                    EcoRI; cDNA made by oligo-dT priming. [illegible]
                    cloned into EcoRI/XhoI [illegible] the following
                    adaptor: [illegible]. [illegible]
                    insert [illegible]. Library constructed [illegible]
                    the laboratory of Gerald [illegible] (University of
                    California, Berkeley) [illegible] cDNA [illegible]
                    (Stratagene) and [illegible]"
BASE COUNT    [illegible]
ORIGIN

Alignment Scores:
Pred. No.:              8.4[illegible]      Length:          [illegible]
Score:                  [illegible]         Matches:         [illegible]
Percent Similarity:     [illegible]         Conservative:    [illegible]
Best Local Similarity:  [illegible]         Mismatches:      [illegible]
Query Match:            [illegible]         Indels:          [illegible]
DB:                     [illegible]         Gaps:            [illegible]

SW[illegible] ([illegible]) x [illegible] ([illegible])

Qy       1 HisGlyTrpSerTyrGlyGlyTyrLeu 9
           [illegible]
Db     [illegible] [illegible] [illegible]

RESULT [illegible]
AL[illegible]
LOCUS       AL[illegible]       [illegible] bp    mRNA    linear   EST [illegible]
DEFINITION  DKFZp434[illegible]723_r1 434 (synonym: htes3) Homo sapiens cDNA [illegible]
            DKFZp434[illegible]723 5', mRNA sequence.
ACCESSION   AL04[illegible]
VERSION     AL04[illegible]  GI:[illegible]
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to [illegible])
  AUTHORS   [illegible], [illegible], Mewes,H.W., [illegible] and Wie[illegible]
  TITLE     EST (Bl[illegible] et al.)
  JOURNAL   Unpublished (1[illegible])
COMMENT     Contact: [illegible]
            MIPS
            Am K[illegible] [illegible] Martinsried, Germany
            This is the 5' sequence of the [illegible]
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); email: s.wiemann@dkfz-heidelberg.de
            sequenced by LMU (Ludwig Maximilians University,
            Munich/Germany) within the cDNA sequencing consortium of the German
            Genome Project.
            No [illegible] sequence available.
            This clone (DKFZp434[illegible]) is available at the RZPD [illegible]
            Please contact the RZPD [illegible]
            Berlin-Charlottenburg, GERMANY [illegible]
FEATURES             Location/Qualifiers
     source          1..[illegible]
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp434[illegible]"
                     /clone_lib="434 (synonym: htes3)"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /lab_host="[illegible]"
                     /note="Vector: pS[illegible] [illegible]"
BASE COUNT    [illegible]
ORIGIN
```

```
Alignment Scores:
Pred. No.:              [illegible]
Score:                  [illegible]
Percent Similarity:     [illegible]
Best Local Similarity:  [illegible]
Query Match:            [illegible]
DB:                     [illegible]

SW[illegible]464 ([illegible]) x [illegible]

Qy       [illegible]

Db       [illegible]

RESULT [illegible]
[illegible]
LOCUS       [illegible]
DEFINITION  [illegible] NIH MGC [illegible] cDNA [illegible]
            cDNA [illegible]
ACCESSION   [illegible]
VERSION     [illegible]
KEYWORDS    [illegible]
SOURCE      [illegible]
  ORGANISM  [illegible]
            [illegible]
            [illegible]
REFERENCE   1  (bases 1 to [illegible])
  AUTHORS   NIH MGC http://mgc[illegible]
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     [illegible]
            [illegible]
FEATURES             Location/Qualifiers
     source          [illegible]
                     /[illegible]
                     /db_xref="[illegible]"
                     /clone="[illegible]"
                     /[illegible]
BASE COUNT    [illegible]
ORIGIN

Alignment Scores:
Pred. No.:              [illegible]
Score:                  [illegible]
Percent Similarity:     [illegible]
Best Local Similarity:  [illegible]
Query Match:            [illegible]
DB:                     [illegible]

SW[illegible]464 [illegible]

Qy       [illegible]

Db       [illegible]

RESULT [illegible]
```

```
                     /db_xref="taxon:[illegible]"
                     /clone="IMAGE:[illegible]"
                     /clone_lib="N[illegible]"
                     /tissue_type="[illegible]"
                     /dev_stage="[illegible]"
                     /lab_host="DH10B"
                     /note="Organ: [illegible] summary: Vector: pCMV-SPORT6; Site_1: [illegible]
                     Site_2: Not I; Cloned [illegible]. Primer: [illegible]
                     library [illegible]
                     [illegible]
                     NIH"
BASE COUNT     [illegible] a    [illegible] c    [illegible] g    [illegible] t
ORIGIN

Alignment Scores:
Pred. No.:                [illegible]          Length:          [illegible]
Score:                    [illegible]          Matches:         [illegible]
Percent Similarity:       [illegible]          Conservative:    [illegible]
Best Local Similarity:    [illegible]          Mismatches:      [illegible]
Query Match:              [illegible]          Indels:          [illegible]
DB:                       [illegible]          Gaps:            [illegible]

SW [illegible] x BE7[illegible] ([illegible])

[illegible]

Db     [illegible]

RESULT [illegible]
BE7[illegible]
LOCUS       BE7[illegible]      [illegible] bp    mRNA    linear   EST [illegible]
DEFINITION  [illegible] NIH_MGC_[illegible] Homo sapiens cDNA clone IMAGE:[illegible]
            mRNA sequence.
ACCESSION   BE78[illegible]
VERSION     BE78[illegible]  GI:[illegible]
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to [illegible])
  AUTHORS   NIH-MGC http://[illegible]
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: [illegible]
            Tissue Procurement: [illegible]
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM[illegible]
            High quality sequence stop: [illegible].
FEATURES             Location/Qualifiers
     source          1..[illegible]
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:[illegible]"
                     /clone_lib="NIH_MGC_[illegible]"
                     /tissue_type="[illegible]"
                     /lab_host="[illegible]"
                     /note="Organ: [illegible]; Vector: pCMV-SPORT6; Site_1: Not I;
                     Site_2: Sal I; Cloned unidirectionally. Primer: [illegible]
                     Average insert size 1.[illegible] kb; library [illegible]
                     [illegible]"
BASE COUNT     [illegible] a    [illegible] c    [illegible] g    [illegible] t
ORIGIN

Alignment Scores:
Pred. No.:                [illegible]          Length:          [illegible]
Score:                    [illegible]          Matches:         [illegible]
```

```
Percent Similarity:       [illegible]
Best Local Similarity:    [illegible]
Query Match:              [illegible]
DB:                       [illegible]

SW [illegible]

[illegible]

Db     [illegible]

RESULT [illegible]
BE7[illegible]
LOCUS       [illegible]
DEFINITION  [illegible]
            mRNA sequence.
ACCESSION   [illegible]
VERSION     [illegible]
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            [illegible]
            [illegible]
REFERENCE   1 [illegible]
  AUTHORS   NIH-MGC [illegible]
  TITLE     National Institutes of Health, [illegible]
  JOURNAL   Unpublished [illegible]
COMMENT     [illegible]
            [illegible]
            cDNA Library [illegible]
            cDNA Library Arrayed by: [illegible]
            DNA Sequencing by: [illegible]
            [illegible]
FEATURES             Location/Qualifiers
     source          [illegible]
                     [illegible]
BASE COUNT     [illegible]
ORIGIN

Alignment Scores:
Pred. No.:                [illegible]
Score:                    [illegible]
Percent Similarity:       [illegible]
Best Local Similarity:    [illegible]
Query Match:              [illegible]
DB:                       [illegible]

SW [illegible]

[illegible]

Db     [illegible]

RESULT [illegible]
BE7[illegible]
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

/organism="Homo sapiens"

/note="Vector: pBluescript ... directionally cloned into the NotI and EcoRI ..."

BASE COUNT

ORIGIN

SW:BE464-1 (1-1) x AA... (1-1)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Xenopus.

JOURNAL Unpublished (1999)

COMMENT

Xenopus tropicalis EST project 2001

FEATURES Location/Qualifiers

/organism="Xenopus tropicalis"

/note="Vector: pCS107 ... cDNA was ... NotI ... EcoRI at the 5' end and NotI at the 3' end."

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.:		Length:	
Score:		Matches:	
Percent Similarity:		Conservative:	
Best Local Similarity:		Mismatches:	
Query Match:		Indels:	
DB:		Gaps:	

SW:BE464-1 (1-1) x ... (1-1)

ACCESSION

VERSION

KEYWORDS EST.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to ...)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

FEATURES Location/Qualifiers

source

/organism="Mus musculus"

```
JOURNAL    [illegible]
MEDLINE    [illegible]
COMMENT    Contact: Chin, H
           National Institute of Mental Health
           [illegible] Executive Blvd. [illegible], Bethesda, M[illegible]
           [illegible], USA
           Tel: 301 443 [illegible]
           Fax: [illegible]
           Email: [illegible]
           The sequence contained in [illegible] that was present [illegible]
           oligonucleotide that was used to prime the synthesis [illegible]
           strand cDNA and therefore this may represent a [illegible] A
           tail. The sequence [illegible] present in the cDNA between the N[illegible]
           and the [illegible] served [illegible]
           amygdala tissue cDNA library [illegible] Source [illegible]
           distribution: Researchers may obtain BMAP cDNA clones from [illegible]
           [illegible]. It should be noted that [illegible]
           small number of additional specialized [illegible]
           cDNAs whose availability will be [illegible] under appropriate [illegible]
           limited collaborative arrangements
           Supplier: MI[illegible] Forward
           [illegible]
FEATURES             [illegible]
     [illegible]          [illegible]
                     /organism="Mus musculus"
                     /strain="[illegible]"
                     /db_xref="taxon:[illegible]"
                     /clone="[illegible]"
                     /clone_lib="NIH_BMAP_M_S4"
                     /dev_stage="[illegible] days"
                     /lab_host="[illegible]"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: NotI; Site_2: EcoRI; The
                     NIH_BMAP_M_S4 library is a subtracted library [illegible]
                     ultimately derived from [illegible]
                     normalized libraries [illegible]
                     (cerebellum, brain stems, olfactory bulbs, hypothalamus,
                     cortex, amygdala, basal ganglia, pineal gland, [illegible]
                     hippocampus) after a series of subtractions to reduce the
                     representation of cDNAs for which ESTs had already been
                     generated. The following serially subtracted libraries
                     were generated in this process: NIH_BMAP_M_S4,
                     NIH_BMAP_M_S[illegible], NIH_BMAP_M_S[illegible], NIH_BMAP_M_S[illegible],
                     NIH_BMAP_M_S[illegible], NIH_BMAP_M_S[illegible]. The subtracted library
                     (NIH_BMAP_M_S4) was constructed as follows: [illegible]
                     cDNA inserts from NIH_BMAP_M_S[illegible], NIH_BMAP_M_S[illegible] [illegible]
                     NIH_BMAP_M_S[illegible] clones for which [illegible] ESTs had been der[illegible]
                     was used as [illegible] driver in a hybridization with [illegible]
                     the NIH_BMAP_M_S[illegible], NIH_BMAP_M_S[illegible], and NIH_BMAP_M_S[illegible]
                     libraries in the form of [illegible] stranded circles. The
                     remaining single-stranded circles (subtracted library)
                     was purified by hydroxyapatite [illegible]
                     converted to double-stranded [illegible] and [illegible]
                     into DH10B bacteria [illegible] to generate the
                     NIH_BMAP_M_S4 library. This procedure has been previous
                     described (Bonaldo, [illegible] and Soares, Genome Research
                     [illegible])
                     [illegible] NIH_BMAP_M_S4
                     TAG_TISSUE amygdala
                     TAG_SEX [illegible]"
BASE COUNT     [illegible]
ORIGIN

Alignment Scores:
Pred. No.:                [illegible]        Length:        [illegible]
Score:                    [illegible]        Matches:       [illegible]
Percent Similarity:       [illegible]        Conservative:  [illegible]
Best Local Similarity:    100.00%            Mismatches:    [illegible]
Query Match:              100.00%            Indels:        [illegible]
DB:                       [illegible]        Gaps:          [illegible]

SW [illegible] x [illegible] (1-4[illegible])
```

```
[illegible]
DE   [illegible]

RESULT [illegible]
AA[illegible]
LOCUS       AA[illegible]
DEFINITION  [illegible] PHOSPHOLIPASE [illegible]
ACCESSION   AA[illegible]
VERSION     AA[illegible]
KEYWORDS    [illegible]
SOURCE      [illegible]
  ORGANISM  [illegible]
REFERENCE   [illegible]
  AUTHORS   [illegible]
  TITLE     [illegible]
  JOURNAL   [illegible]
COMMENT     [illegible]
FEATURES             [illegible]
BASE COUNT     [illegible]
ORIGIN

Alignment Scores:
Pred. No.:                [illegible]
Score:                    [illegible]
Percent Similarity:       [illegible]
Best Local Similarity:    [illegible]
Query Match:              [illegible]
DB:                       [illegible]

SW [illegible]

QY [illegible]

DB [illegible]

RESULT [illegible]
AA[illegible]
LOCUS       AA[illegible]
DEFINITION  [illegible] PHOSPHOLIPASE [illegible]
```

SUMMARIES

look

>FD

ALIGNMENTS

DEFINITION

VERSION

REFERENCE

[illegible]

RESULT [illegible]
ABK[illegible]
XX
[illegible]
XX
[illegible]
XX
[illegible]
XX
KW Human; serine protease; dipeptidyl peptidase IV related protein; [illegible]
KW [illegible] infection; human immunodeficiency virus; HIV [illegible]
KW diabetes; infertility; [illegible]
KW heart failure; hypertension; urinary retention; osteoporosis; [illegible]
KW [illegible] allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; [illegible]
KW [illegible]
XX
[illegible]
XX
[illegible]
XX
[illegible]
XX
[illegible]
XX
PA (FERR-) FERRING BV.
XX
[illegible]
XX
DR WPI; [illegible]/47.
[illegible]
XX
PT New dipeptidyl peptidase IV related [illegible]
PT the proteins, useful for treating [illegible]
PT viral infections, cancers, allergies, neurological disorders, [illegible]
XX
[illegible]
XX
[illegible] The present invention relates to [illegible]
[illegible] proteins (DPRP). The dipeptidyl peptidase IV related proteins [illegible]
[illegible] and nucleic acids encoding them are useful for treating infections
[illegible] such as fungal, bacterial, protozoan and viral infections, particularly
[illegible] infections caused by human immunodeficiency virus (HIV-1 or HIV-2),
[illegible] diabetes [illegible] infertility [illegible]
[illegible] Parkinson's disease, acute heart failure, hypotension,
[illegible] hypertension, urinary retention, osteoporosis, angina pectoris,
[illegible] stroke, ulcers, asthma, allergies, cancers [illegible]
[illegible] psychotic and neurological disorders [illegible]
[illegible] and dyskinesias. These may also be used [illegible]
[illegible] therapeutic agents for the treatment of [illegible] inflammatory and
[illegible] [illegible] disorders. ABK[illegible] human DPRP proteins.
XX
[illegible]

Alignment Scores
[illegible]
Best Local Similarity: 100.0%
[illegible] 24 [illegible]

```
XX

Alignment Scores:

SW-PETCG I (1-9) x AAH... (1-11)

AC   AAH...
XX
DT   26 MAR 2002 (first entry)
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   therapeutic; gene.
XX
PN   WO...
XX
PA   (PEKI-) ... NA
XX
PI   Venter JC, Adams M, Li PWL, Myers EW;
XX
PT   New isolated nucleic acid detection reagent for detecting ... more
PT   genes from Drosophila and for elucidating cell signalling and ...
PT   interaction.
XX
PS   Claim ...; ...pp; Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell ... interaction in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses ... DNA sequences (ABL..., ABL...), expressed DNA
CC   sequences (ABL... ABL...) and the encoded proteins
CC   (AAB... AAB...).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence ...

Alignment Scores:
Pred. No.:                       Length:
Score:                           Matches:
Percent Similarity:              Conservative:
Best Local Similarity:           Mismatches:
Query Match:                     Indels:
DB:                              Gaps:
```

```
Alignment Scores:
Pred. No.:                       Length:
Score:                           Matches:
Percent Similarity:              Conservative:
Best Local Similarity:           Mismatches:
Query Match:                     Indels:
DB:                              Gaps:

SW-PETCG I (1-9) x ABL... (1-11)

DB   ...

ABL...
ABL...
ID   ABL... standard; DNA; ... BP.
XX
AC   ABL...
XX
DT   26 MAR 2002 (first entry)
XX
DE   Drosophila ... expressed ... SEQ ID NO ...
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene.
XX
OS   Drosophila melanogaster.
XX
PN   WO...
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; ...
PR   11-JUL-2000; ...
XX
PA   (PEKI-) ... NA
XX
PI   Venter JC, Adams M, Li PWL, Myers EW;
XX
DR   WPI; ...
DR   P-PSDB; ABP...
XX
PT   New isolated nucleic acid detection reagent for detecting ... more
PT   genes from Drosophila and for elucidating cell signalling and ...
PT   interaction.
XX
PS   Claim ...; ...pp; Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell ... interaction in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses ... DNA sequences (ABL..., ABL...), expressed DNA
CC   sequences (ABL... ABL...) and the encoded proteins
CC   (AAB... AAB...).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence ...

Alignment Scores:
Pred. No.:                       Length:
Score:                           Matches:
Percent Similarity:              Conservative:
Best Local Similarity:           Mismatches:
Query Match:                     Indels:
DB:                              Gaps:
```

```
SQ   Sequence [illegible]

Alignment Scores:
Pred. No.:                 [illegible]     Length:        [illegible]
Score:                     [illegible]     Matches:       [illegible]
Percent Similarity:        [illegible]     Conservative:  [illegible]
Best Local Similarity:     100.00%         Mismatches:    [illegible]
Query Match:               100.00%         Indels:        [illegible]
DB:                        [illegible]     Gaps:          [illegible]

SW [illegible] x AA785094 [illegible]

Qy   [illegible]

Db   [illegible] TATGAGATA [illegible]

RESULT [illegible]
AA[illegible]
ID   AA[illegible] standard; cDNA; [illegible]
XX
AC   AA[illegible]
XX
DT   [illegible] (first entry)
XX
DE   Human dipeptidyl peptidase 8 (DPP8) cDNA.
XX
KW   Human; dipeptidyl peptidase; DPP; neoplasia; type II diabetes; [illegible]
KW   autoimmunity; human immune deficiency virus; HIV infection; [illegible]
KW   graft rejection; antidiabetic; antiinflammatory; immunosuppressive;
KW   antiviral; enzyme; detection.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             [illegible]
FT                   /*tag= a
FT                   /product= "Human DPP8 protein"
XX
PN   WO[illegible] A1
XX
PD   [illegible] MAY 2002
XX
PF   [illegible] WO[illegible] AU[illegible]
XX
PR   [illegible] 2000; [illegible] PAT[illegible]
XX
PA   (UNSY ) UNIV SYDNEY.
XX
PI   Abbott CA, Gorrell MD;
XX
DR   WPI; 2002-[illegible]
DR   P-PSDB; AAB[illegible]
XX
PT   New dipeptidyl peptidase (DPP) peptides, useful for screening
PT   inhibitors of DPP catalytic activity, which may be employed for treating
PT   neoplasia, type II diabetes, [illegible] graft
PT   rejection and HIV infection -
XX
PS   Example; [illegible] English.
XX
CC   The present invention relates to dipeptidyl peptidase (DPP) [illegible]
CC   polynucleotides encoding such proteins. The DPP peptides are useful for
CC   screening inhibitors of DPP catalytic activity. The inhibitors are useful
CC   for treating neoplasia, type II diabetes, [illegible]
CC   rejection and HIV (human immuno deficiency virus) [illegible] The present
CC   sequence is human DPP8 cDNA.
XX
SQ   Sequence [illegible]

Alignment Scores:
Pred. No.:                 [illegible]     Length:        [illegible]
Score:                     [illegible]     Matches:       [illegible]
Percent Similarity:        [illegible]     Conservative:  [illegible]

Best Local Similarity:     [illegible]
Query Match:               [illegible]
DB:                        [illegible]

SW [illegible]

Qy   [illegible]

Db   [illegible]

RESULT [illegible]
AA[illegible]
ID   AA[illegible]
XX
AC   AA[illegible]
XX
DT   [illegible]
XX
DE   [illegible] DNA [illegible]
XX
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
XX
OS   [illegible]
XX
FH   [illegible]
FT   [illegible]
FT   [illegible]
FT   [illegible]
FT   [illegible]
FT   [illegible]
XX
PN   W[illegible] A[illegible]
XX
PD   [illegible]
XX
PF   [illegible]
XX
PR   [illegible]
XX
PA   [illegible]
XX
PI   [illegible]
XX
DR   W[illegible]
DR   [illegible] AA[illegible]
XX
PT   [illegible]
PT   [illegible]
PT   [illegible]
XX
PS   [illegible]
XX
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
```

[illegible]

XX
KW Human; [illegible] proteinase [illegible]
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
[illegible]

XX
[illegible]

XX
KW Human; dipeptidyl aminopeptidase; DPP8; prolyl oligopeptidase; [illegible]
KW [illegible] diabetes mellitus; glucose intolerance; [illegible]
KW immunosuppression; [illegible]
XX
FT [illegible] "Human DPP8"
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PT New human dipeptidyl aminopeptidase (DPP8) useful for [illegible]
XX
CC This sequence [illegible] human dipeptidyl aminopeptidase (DPP8). DPP8
CC has substrate specificity for H-Ala-Pro-pNA, H-Gly-Pro-pNA and
CC H-Arg-Pro-pNA [illegible] because it is capable of hydrolysing the
CC peptide bond [illegible] to Pro in each of these compounds. DPP8
CC [illegible] DPP8 is useful for [illegible]
CC [illegible] activated T cell which [illegible]
CC [illegible] DPP8 gene expression in a T cell. The [illegible]
CC of DPP8 expression [illegible] detected by detecting the amount of DPP8 [illegible]
CC [illegible] useful for identifying a molecule [illegible]
CC [illegible] of the substrate by DPP8. Molecules
CC identified as [illegible] DPP8 catalytic activity may be useful for
CC treating [illegible] deficiency, lowering [illegible]
CC [illegible] diabetes mellitus and [illegible]
CC [illegible] enhancing [illegible]
CC as immunosuppressants
XX

```
XX
KW   [illegible] protein; [illegible] peptidase; [illegible] therapy; [illegible] peptidase;
KW   cancer; cardiovascular disease; autoimmune disease; [illegible]
KW   neuronal disorder; vascular disorder; prostate disorder; cytostatic;
KW   antidiabetic; antiarthritic; antiasthmatic; anti-inflammatory;
KW   diabetes mellitus; arthritis; multiple sclerosis; asthma;
KW   Crohn's disease; neuronal disorder; [illegible] disease; ss.
XX
OS   Homo sapiens.
XX
PN   W[illegible]47[illegible]-A2.
XX
PD   [illegible]
XX
PF   [illegible] 2001; 2001WO-US[illegible].
XX
PR   [illegible] 2000; 2000US-[illegible]P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Meyers RA,  Williamson M;
XX
DR   WPI; 2002-[illegible]/[illegible].
DR   P-PSDB; AAG[illegible]41[illegible].
XX
PT   New polypeptides [illegible] human [illegible] peptidase [illegible]
PT   useful as diagnostic targets and therapeutic agents [illegible]
PT   cancer, lymphoma and leukemia [illegible]
XX
PS   Claim [illegible]; Page [illegible]; [illegible]pp; English.
XX
CC   This invention relates to an isolated [illegible] human protein
CC   [illegible] peptidase, which is cytostatic, antidiabetic, antiarthritic,
CC   neuroprotective, antithyroid, dermatological, antipsoriatic,
CC   antiasthmatic, ophthalmological, anti-inflammatory, [illegible],
CC   antiparkinsonian, anticonvulsant, [illegible], [illegible],
CC   antianginal, cardiant, [illegible] and
CC   [illegible] in its activity. Uses include gene therapy, expression [illegible]
CC   activity of [illegible] protein modulators. It is useful for identifying
CC   [illegible] and which binds [illegible] and can be used in preventing, treating
CC   or detecting a cellular proliferative or differentiative disorders.
CC   The [illegible] molecules can act as novel diagnostic targets and therapeutic
CC   agents for controlling disorders associated with the aberrant [illegible]
CC   or degradation of peptide hormones [illegible] associated with [illegible]
CC   differentiation and proliferation such as cancer, immune function,
CC   reproductive, neurological, and cardiovascular function. The [illegible]
CC   molecules are thus useful for treating and preventing cellular
CC   proliferative and differentiative disorders [illegible]
CC   disorders, immune disorders such as autoimmune diseases, diabetes
CC   mellitus, arthritis, multiple sclerosis, asthma, Crohn's disease,
CC   neuronal disorders, [illegible] diseases, [illegible] disorders and
CC   [illegible] or pain disorders. This sequence represents the DNA
CC   encoding sequence of [illegible] human [illegible] peptidase. This
CC   sequence represents the [illegible] sequence [illegible]
CC   in between the start and the stop codon [illegible] sequence represents
CC   AA[illegible].
XX
SQ   Sequence 2[illegible] BP; [illegible] A; [illegible] C; [illegible] G; [illegible] T; [illegible] other;

Alignment Scores:
Pred. No.:              1.2[illegible]        Length:           [illegible]
Score:                  [illegible]           Matches:          [illegible]
Percent Similarity:     1[illegible]%         Conservative:     [illegible]
Best Local Similarity:  1[illegible]%         Mismatches:       [illegible]
Query Match:            1[illegible]%         Indels:           [illegible]
DB:                     2[illegible]          Gaps:             [illegible]

SW[illegible] ([illegible]) x AA[illegible] ([illegible])

Qy        [illegible] HisTyrTrpSerTyrGlyGlyTyrLeu[illegible]
              |||||||||||||
Db       [illegible] CA[illegible]GTATGGAGGATATC[illegible] 2229
```

```
XX
AC   [illegible]
XX
DT   [illegible]
XX
DE   [illegible]
XX
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
XX
OS   Homo sapiens.
XX
PN   W[illegible]
XX
PD   [illegible]
XX
PF   [illegible]
XX
PR   [illegible]
XX
PA   [illegible] INC.
XX
PI   [illegible]
XX
DR   WPI; 2002-[illegible].
DR   P-PSDB; AB[illegible].
XX
PT   [illegible]
PT   [illegible]
PT   [illegible]
PT
XX
PS   [illegible]
XX
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
XX
SQ   [illegible]

Alignment Scores:
Pred. No.:              [illegible]
Score:                  [illegible]
Percent Similarity:     [illegible]
Best Local Similarity:  [illegible]
Query Match:            [illegible]
DB:                     [illegible]

SW[illegible]

Qy        [illegible]

Db       [illegible]
```

FT /product= "Human DPP8 [illegible]"
FT /note= "No stop [illegible]"
XX
PN WO[illegible]
XX
PA [illegible] UNIV SYDNEY
XX
PI Abbott CA, [illegible] MD;
XX
DR WPI; [illegible]
DR P-PSDB; AAB[illegible]
XX
PT New human dipeptidyl aminopeptidase (DPP8) [illegible]
XX
PS [illegible]
XX
CC [illegible] dipeptidyl aminopeptidase (DPP8). [illegible] H-Ala-Pro-pNA, H-[illegible]-pNA and H-Arg-Pro-pNA. [illegible] prolyl oligopeptidase and a dipeptidyl peptidase, because it is capable of hydrolysing the peptide bond C-terminal to Pro in each of these compounds. DPP8 is homologous with human DPPIV. DPP8 is useful for cleaving a substrate, and for detecting an activated T cell which involves measuring the level of DPP8 gene expression in a T cell. The level of DPP8 expression is detected by detecting the amount of DPP8 RNA in the cell. It is also useful for identifying a molecule capable of inhibiting the cleavage of the substrate by DPP8. Molecules identified as inhibiting DPP8 catalytic activity may be used for [illegible]
XX
[illegible]

[illegible]
XX
AC AAB[illegible];
XX
[illegible]
XX
DE [illegible]
XX
KW [illegible]

KW Human [illegible]
XX
OS Homo sapiens.
XX
FH Key [illegible]
FT [illegible]
FT /product= "Human DPP8 [illegible]"
XX
PN WO[illegible]
XX
PD [illegible]
XX
PF [illegible]
XX
PR [illegible]
PR [illegible]
XX
PA (UNSY) UNIV SYDNEY.
XX
PI Abbott CA, [illegible]
XX
DR WPI; [illegible]
DR P-PSDB; AAB[illegible]
XX
PT New human dipeptidyl [illegible] peptidase (DPP8) useful [illegible] substrates, [illegible] DPP8 catalytic activity, [illegible] and for detecting activated T cells
XX
PS Claim [illegible]; [illegible]pp; English.
XX
CC The sequences [illegible] AAB[illegible]-97 encode fragments of human dipeptidyl aminopeptidase (DPP8). DPP8 has substrate specificity for H-Ala-Pro-pNA, H-[illegible]-Pro-pNA and H-Arg-Pro-pNA. Therefore, it is a prolyl oligopeptidase and a dipeptidyl peptidase, because it is capable of hydrolysing the peptide bond C-terminal to Pro in each of these compounds. DPP8 is homologous with human DPPIV. DPP8 is useful for cleaving a substrate, and for detecting an activated T cell which involves measuring the level of DPP8 gene expression in a T cell. The level of DPP8 expression is detected by detecting the amount of DPP8 RNA in the cell. It is also useful for identifying a molecule capable of inhibiting the cleavage of the substrate by DPP8. Molecules identified as inhibiting DPP8 catalytic activity may be used for treating [illegible] deficiency, lowering [illegible] diabetes mellitus and other disorders [illegible], enhancing mucosal regeneration and as immunosuppressants.
XX
SQ Sequence [illegible] AA; [illegible] C; [illegible] G; [illegible] T; 0 other;

[illegible]

RESULT [illegible]
AAB[illegible]
ID AAB[illegible] standard; [illegible]
XX
AC AAB[illegible];
XX
[illegible]
XX
DE [illegible] peptidase

```
XX
PI   [illegible] Tang YT, Baughn [illegible], Patterson MR, Azimzai Y, [illegible]
PI   [illegible]
XX
DR   WPI; [illegible]/A.
DR   P-PSDB; AAB[illegible].
XX
PT   New protease (inhibitors) useful for diagnosis and treatment of
PT   autoimmune/inflammatory disorders such as acquired immunodeficiency
PT   syndrome, Cushing's disease, Addison's disease and cell proliferative
PT   disorders such as cancer
XX
PS   Claim 7; Page 121; 134pp; English.
XX
CC   AAF[illegible] - AAF[illegible] encode the human proteases and protease inhibitors
CC   (PRIMs) given in AAB74668 to AAB74694. The PRIMs can have activity such
CC   as: anti-human immunodeficiency virus (HIV) [illegible]
CC   immunostimulant; immunomodulator; anti-inflammatory; immunosuppressive;
CC   nephrotropic; antigout; thyromimetic; cytostatic; antibacterial;
CC   fungicide; protozoacide; antiarteriosclerotic; antiatherosclerotic;
CC   virucide; antipsoriatic; and hepatotropic. PRIM polynucleotide and
CC   protein sequences can be used in the diagnosis, treatment and prevention
CC   of autoimmune/inflammatory disorders such as AIDS, Addison's syndrome,
CC   severe combined immunodeficiency disease (SCID), Chediak-Higashi
CC   syndrome, Cushing's disease, Addison's disease, autoimmune thyroiditis,
CC   Crohn's disease, diabetes mellitus, Goodpasture's syndrome, [illegible]
CC   Grave's diseases, Hashimoto's thyroiditis, Sjogren's syndrome, Werner's
CC   syndrome, viral, bacterial, fungal, parasitic, protozoal, and helminthic
CC   infections and cell proliferative disorder such as arteriosclerosis,
CC   atherosclerosis, cirrhosis, hepatitis, psoriasis and cancer. PRIM
CC   polynucleotide sequences can be used in somatic or germline gene therapy
CC   and in diagnosis of diseases. They can also be used in generating
CC   hybridisation probes useful in mapping [illegible]
CC   sequences and in molecular biology techniques.
XX
SQ   Sequence [illegible]

Alignment Scores:
Pred. No.:               [illegible]      Length:          [illegible]
Score:                   [illegible]      Matches:         [illegible]
Percent Similarity:      100.0%           Conservative:    [illegible]
Best Local Similarity:   100.0%           Mismatches:      [illegible]
Query Match:             [illegible]      Indels:          [illegible]
DB:                      22               Gaps:            [illegible]

SWISSPROT [illegible] x AAF[illegible] ([illegible])

Qy   [illegible] HisGlyIleSerPheGlyCysLys [illegible]
Db   [illegible] CATGGTATCCTATGGAGAATACCT [illegible]

RESULT [illegible]
AAF[illegible]
ID   AAF[illegible] standard; cDNA; [illegible]
XX
AC   AAF[illegible]
XX
DT   [illegible] MAY [illegible] (first entry)
XX
DE   Human polynucleotide SEQ ID NO [illegible].
XX
KW   Cytostatic; immunosuppressive; [illegible]
KW   anti-inflammatory; hepatotropic; antidiabetic; [illegible]
KW   veterinary; anticonvulsant; antibacterial; [illegible]
KW   [illegible] gene therapy; [illegible]
KW   [illegible] disease; infection; human; [illegible]
XX
OS   Homo sapiens.
XX
PN   WO200[illegible]-A[illegible].
XX
PD   [illegible] NOV-200[illegible].
XX
```

```
PF   [illegible] MAY [illegible]
XX
PR   [illegible] MAY [illegible]
XX
PR   [illegible] MAY [illegible] MAY [illegible]
XX
PA   [illegible]
XX
DR   WPI; [illegible]
DR   P-PSDB; AAB[illegible]
XX
PT   [illegible]
PT   [illegible]
PT   [illegible]
PT   [illegible]
XX
PS   [illegible]
XX
CC   The invention [illegible]
CC   (ABB[illegible] - ABB[illegible]) [illegible]
CC   medical conditions [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   disease [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
XX
SQ   Sequence [illegible]

Alignment Scores:
Pred. No.:               [illegible]
Score:                   [illegible]
Percent Similarity:      [illegible]
Best Local Similarity:   [illegible]
Query Match:             [illegible]
DB:                      [illegible]

SWISSPROT [illegible]

Qy   [illegible] HisGlyIleSer [illegible]
Db   [illegible]

RESULT [illegible]
AAF[illegible]
ID   AAF[illegible] standard; cDNA; [illegible]
XX
AC   AAF[illegible]
XX
DT   [illegible] JUN [illegible] (first entry)
XX
DE   [illegible]
XX
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
XX
OS   Homo sapiens.
XX
PN   [illegible]
PD   [illegible]
```

XX
[illegible] used in prevention, treatment [illegible] medical condition [illegible]
XX
[illegible] Claim 1; SEQ ID NO 8[illegible]; 120pp; [illegible] Sequence Listing: English.
XX
The present invention provides the protein and coding sequences [illegible] number of human reproductive system related antigens. These can be used [illegible] the prevention and treatment [illegible] reproductive system [illegible] [illegible] cancer. The present sequence is [illegible]
XX
[illegible] Sequence [illegible] BP; [illegible] A; [illegible] other;

Alignment Scores:
[illegible]

[illegible] AAI[illegible] ([illegible])

[illegible]

XX
[illegible] AAB[illegible] standard; cDNA; [illegible] BP.
XX
AC AAB[illegible];
XX
[illegible] (first entry)
XX
[illegible] protease [illegible]
XX
KW [illegible] protease [illegible]
KW [illegible] protease mediated disorder; proliferative disorder;
KW differentiative disorder; developmental disorder;
KW haematopoietic disorder;
XX
[illegible] sapiens.
XX
[illegible]
XX
[illegible]
XX
[illegible] MAR[illegible]
XX
[illegible] MAR[illegible]
XX
[illegible] MILLENNIUM PHARM INC.
XX
[illegible]
XX
[illegible]
XX
New [illegible] protease homolog [illegible] protease [illegible] useful in identifying agonists and antagonists [illegible] treatment of protease mediated disorders.
XX
[illegible]
XX
[illegible] isolated polypeptide encoded by the nucleic acid [illegible] the nucleic acid into a host cell and culturing under conditions to express the protein from the nucleic acid [illegible] detect the encoded protein [illegible] [illegible] identifying agents that [illegible]

[illegible] that modulates the expression [illegible] acid [illegible] (e.g. gene therapy). The nucleic acid can be used to identify [illegible] that modulates the expression or activity of the nucleic acid and can be used to isolate the protein. The nucleic acid can be used in diagnostic assays for determining nucleic acid expression as well as activity in the context of a biological sample (e.g. [illegible] serum [illegible]) to determine whether an individual has a disease or disorder [illegible] risk of developing a disease or disorder [illegible] with aberrant [illegible] activity of the nucleic acid. The nucleic acid can be used to detect mutations in protease genes and [illegible] expression [illegible] such as mRNA. The nucleic acid can be [illegible] hybridisation [illegible] to detect naturally occurring genetic [illegible] a protease gene. The nucleic acid can be used in drug screening [illegible] to identify agonists and antagonists that can be used to diagnose and treat such protease mediated disorders e.g. proliferative, differentiative [illegible] developmental or haematopoietic disorders. [illegible] acid can be used [illegible] probes, primers, in [illegible] assays [illegible] patterns of gene expression [illegible] disease [illegible] and [illegible] treatment [illegible] present sequence represents [illegible] cDNA sequence [illegible] protease cDNA sequence.
XX
SQ Sequence [illegible] other;

Alignment Scores:
[illegible]

[illegible] AAB[illegible] ([illegible])

[illegible]

[illegible]
AAB[illegible]
ID AAB[illegible] standard; cDNA; [illegible] BP.
XX
AC AAB[illegible];
XX
[illegible]
XX
[illegible] Human protease [illegible] cDNA.
XX
KW Human [illegible] protease [illegible]
KW [illegible] cytostatic [illegible] immunomodulator; antiinflammatory [illegible]
KW antithyroid [illegible] immunosuppressive; nephrotropic; antianemic; [illegible]
KW [illegible] antibacterial; fungicide; protozoacide; antiinfective [illegible]
KW antifibrotic [illegible] antipsoriatic; virucide; hepatotropic; [illegible]
KW [illegible] inflammatory disease; AIDS [illegible]
KW [illegible] disease [illegible]
KW [illegible]
KW [illegible]
KW [illegible] proliferative disorder; arthritis [illegible]
KW atherosclerosis; cirrhosis; hepatitis; psoriasis; ss.
XX
[illegible] Homo sapiens.
XX
[illegible] WO[illegible]
XX
[illegible]
XX
[illegible]
XX
[illegible]
[illegible]
XX
PA (INCY-) INCYTE GENOMICS INC.

SUMMARIES

Result No.

ALIGNMENTS

RESULT 1

XX

XX

XX

XX

KW

KW

XX

XX

XX

```
;  APPLICANT: [illegible]
;  APPLICANT: Akinsanya, Karen
;  APPLICANT: [illegible], Pierre
;  APPLICANT: [illegible]
;  TITLE OF INVENTION: NOVEL SERINE [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: US/[illegible]
;  CURRENT FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: PatentIn [illegible]
; SEQ ID NO [illegible]
;   LENGTH: [illegible]
;   TYPE: DNA
;   ORGANISM: Homo sapiens
[illegible]

Alignment Scores:
[illegible]                    [illegible]        [illegible]     [illegible]
Score:                         [illegible]        Matches:        [illegible]
Percent Similarity:            [illegible]        Conservative:   [illegible]
Best Local Similarity:         [illegible]        Mismatches:     [illegible]
Query Match:                   [illegible]        Indels:         [illegible]
DB:                            [illegible]        Gaps:           [illegible]

SW [illegible] (1-[illegible]) x US-[illegible] ([illegible])

Qy        [illegible]AsnValHis[illegible]
              [illegible]
Db        [illegible]AAAA[illegible]

Search completed: February [illegible]
Job time : [illegible] secs
```

; GENERAL INFORMATION:

Alignment Scores:

[illegible]

```
Query Match:              100.0%       Indels:         0
DB:                       10           Gaps:           0

SWOPE464-2 (1-9) x US-09-976-674-2 (1-2671)

Qy          1 LeuAspIleAsnValHisPheAlaHis 9
              |||||||||||||||||||||||||||
Db       2453 CTGGATATAAATGTCCATTTTGCACAT 2479

RESULT 2
US-09-976-674-14
; Sequence 14, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
;  APPLICANT: Qi, Steve
;  APPLICANT:  Akinsanya, Karen
;  APPLICANT:  Riviere, Pierre
;  APPLICANT:  Junien, Jean-Louis
;  TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
;  FILE REFERENCE: 70669
;  CURRENT APPLICATION NUMBER: US/09/976,674
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: US 60/240,117
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
;   LENGTH: 4309
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-976-674-14

Alignment Scores:
Pred. No.:                0.205        Length:         4309
Score:                    51.00        Matches:        9
Percent Similarity:       100.00%      Conservative:   0
Best Local Similarity:    100.00%      Mismatches:     0
Query Match:              100.0%       Indels:         0
DB:                       10           Gaps:           0

SWOPE464-2 (1-9) x US-09-976-674-14 (1-4309)

Qy          1 LeuAspIleAsnValHisPheAlaHis 9
              |||||||||||||||||||||||||||
Db       2141 CTGGATATAAATGTCCATTTTGCACAT 2167

RESULT 3
US-09-976-674-8
; Sequence 8, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
;  APPLICANT: Qi, Steve
;  APPLICANT:  Akinsanya, Karen
;  APPLICANT:  Riviere, Pierre
;  APPLICANT:  Junien, Jean-Louis
;  TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
;  FILE REFERENCE: 70669
;  CURRENT APPLICATION NUMBER: US/09/976,674
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: US 60/240,117
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 4523
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-976-674-8

Alignment Scores:
Pred. No.:                0.746        Length:         4523
Score:                    51.00        Matches:        9
Percent Similarity:       100.00%      Conservative:   0
```

```
Best Local Similarity:    [illegible]
Query Match:              [illegible]
DB:                       10           [illegible]

SWOPE464-2 (1-9) x [illegible]

Qy          1 LeuAspIleAsnValHis[illegible]
              [illegible]
Db       [illegible]

RESULT 4
US-09-976-674-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. US2002[illegible]
; GENERAL INFORMATION:
;  APPLICANT: Qi, Steve
;  APPLICANT:  Akinsanya, Karen
;  APPLICANT:  Riviere, Pierre
;  APPLICANT:  Junien, Jean-Louis
;  TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
;  FILE REFERENCE: 70669
;  CURRENT APPLICATION NUMBER: [illegible]
;  CURRENT FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: PatentIn version [illegible]
; SEQ ID NO [illegible]
;   LENGTH: [illegible]
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-976-674-[illegible]

Alignment Scores:
Pred. No.:                [illegible]
Score:                    [illegible]
Percent Similarity:       [illegible]
Best Local Similarity:    [illegible]
Query Match:              [illegible]
DB:                       10           [illegible]

SWOPE464-2 (1-9) x [illegible]

Qy          1 LeuAspIleAsnVal[illegible]
              [illegible]
Db       [illegible]

RESULT 5
US-09-976-674-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. US2002[illegible]
; GENERAL INFORMATION:
;  APPLICANT: Qi, Steve
;  APPLICANT:  Akinsanya, Karen
;  APPLICANT:  Riviere, Pierre
;  APPLICANT:  Junien, Jean-Louis
;  TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
;  FILE REFERENCE: 70669
;  CURRENT APPLICATION NUMBER: [illegible]
;  CURRENT FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: PatentIn version [illegible]
; SEQ ID NO [illegible]
;   LENGTH: [illegible]
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-976-674-[illegible]

Alignment Scores:
Pred. No.:                [illegible]
Score:                    [illegible]
```

DB Published Applications NA [illegible] MINMATCH

[illegible] START 1 END 1 MATRIX

SUMMARIES

ALIGNMENTS

SFD

```
Alignment Scores:
Pred. No.:            [illegible]      Length:      [illegible]
Score:                [illegible]      Matches:     [illegible]
Percent Similarity:   [illegible]      Conservative: [illegible]
Best Local Similarity: [illegible]     Mismatches:  [illegible]
Query Match:          [illegible]      Indels:      [illegible]
DB:                   [illegible]      Gaps:        [illegible]

SWOPE464-1 (1-9) x US [illegible]

Qy    1 HisGlyTrpSerTyrGlyTyrLys[illegible]
Db    [illegible]

RESULT [illegible]
US [illegible]
; Sequence [illegible], Application US/ [illegible]
; Patent No. US2002[illegible]1584[illegible]A1
; GENERAL INFORMATION:
;  APPLICANT: Qi, Steve
;  APPLICANT: Akinsanya, Karen
;  APPLICANT: Riviere, Pierre
;  APPLICANT: [illegible], Jean-Louis
;  TITLE OF INVENTION: NOVEL SERINE PROTEASE [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: US/[illegible]
;  CURRENT FILING DATE: 2001-[illegible]
;  PRIOR APPLICATION NUMBER: US 60/240,117
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: PatentIn version [illegible]
; SEQ ID NO [illegible]
;  LENGTH: [illegible]
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US [illegible]

Alignment Scores:
Pred. No.:            [illegible]      Length:      [illegible]
Score:                [illegible]      Matches:     [illegible]
Percent Similarity:   [illegible]      Conservative: [illegible]
Best Local Similarity: [illegible]     Mismatches:  [illegible]
Query Match:          [illegible]      Indels:      [illegible]
DB:                   [illegible]      Gaps:        [illegible]

SWOPE464-1 (1-9) x US [illegible]

Qy    1 HisGlyTrpSerTyrGlyTyrLys[illegible]
Db    [illegible]

RESULT [illegible]
US [illegible]
; Sequence [illegible], Application US/ [illegible]
; Patent No. US2002[illegible]A1
; GENERAL INFORMATION:
;  APPLICANT: Qi, Steve
;  APPLICANT: Akinsanya, Karen
;  APPLICANT: Riviere, Pierre
;  APPLICANT: [illegible], Jean-Louis
;  TITLE OF INVENTION: NOVEL SERINE PROTEASE [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: US/[illegible]
;  CURRENT FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: US [illegible]
;  PRIOR FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: PatentIn version [illegible]
; SEQ ID NO [illegible]
;  LENGTH: [illegible]
;  TYPE: DNA
```

[illegible]

[illegible]

Alignment Scores:
Matches:
Indels:
Gaps:

[illegible]

; GENERAL INFORMATION
; APPLICANT:
; TITLE OF INVENTION:
; CURRENT APPLICATION NUMBER:
; CURRENT FILING DATE:
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS:
; SOFTWARE:
; SEQ ID NO
; LENGTH:
; TYPE: DNA
; ORGANISM: Homo sapiens

Alignment Scores:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:

[illegible]

; Sequence [illegible], Application US/[illegible]
; Publication No. [illegible]
; GENERAL INFORMATION
; APPLICANT: [illegible]
; TITLE OF INVENTION: [illegible] Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: [illegible]
; CURRENT APPLICATION NUMBER: [illegible]

[illegible]

; CURRENT FILING DATE:
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS:
; SOFTWARE: [illegible] Version [illegible]
; SEQ ID NO
; LENGTH:
; TYPE: DNA
; ORGANISM: [illegible]
; FEATURE:
; NAME/KEY:
; LOCATION:

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:

[illegible]

RESULT [illegible]
; Sequence [illegible], Application US/[illegible]
; Publication No. [illegible]
; GENERAL INFORMATION
; APPLICANT: [illegible]
; TITLE OF INVENTION: [illegible] Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER:
; CURRENT FILING DATE:
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; PRIOR APPLICATION NUMBER: 09/[illegible]
; PRIOR FILING DATE:
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS:
; SOFTWARE: [illegible] Version [illegible]
; SEQ ID NO
; LENGTH:
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY:
; LOCATION:

```
Query Match:            [illegible]      Indels:
Db:                     [illegible]      Gaps:

SW [illegible] x [illegible]

Qy       [illegible]

Db       [illegible]

RESULT 2
US-09-[illegible]-674-14
; Sequence 14, Application US/[illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;  APPLICANT: Li, Steve
;  APPLICANT: Akinsanya, Karen
;  APPLICANT: Riviere, Pierre
;  APPLICANT: Junien, Jean-Louis
;  TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
;  FILE REFERENCE: 7066[illegible]
;  CURRENT APPLICATION NUMBER: US/09/[illegible]
;  CURRENT FILING DATE: 2001-[illegible]-12
;  PRIOR APPLICATION NUMBER: US 60/24[illegible]
;  PRIOR FILING DATE: 2000-[illegible]-12
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: PatentIn version [illegible]
; SEQ ID NO 14
;   LENGTH: [illegible]
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-[illegible]-14

Alignment Scores:
Pred. No.:              [illegible]      Length:        [illegible]
Score:                  59.[illegible]   Matches:       [illegible]
Percent Similarity:     [illegible]      Conservative:  [illegible]
Best Local Similarity:  [illegible]      Mismatches:
Query Match:            [illegible]      Indels:
DB:                     [illegible]      Gaps:

SWOPE464 1 (1-9) x US-09-976-674-14 (1-4[illegible])

Qy       1 HisGlyTrpSerTyrGlyGlyTyrLeu [illegible]

Db       [illegible] CACGGCTGGTCCTATGGAGGATACCTC [illegible]

RESULT [illegible]
US-[illegible]-2
; Sequence 2, Application US/[illegible]
; Patent No. US2002[illegible]584A1
; GENERAL INFORMATION:
;  APPLICANT: Li, Steve
;  APPLICANT: Akinsanya, Karen
;  APPLICANT: Riviere, Pierre
;  APPLICANT: Junien, Jean Louis
;  TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: US/09/[illegible]
;  CURRENT FILING DATE: 2001-[illegible]-12
;  PRIOR APPLICATION NUMBER: US 60/24[illegible]
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: PatentIn version [illegible]
; SEQ ID NO 2
;   LENGTH: 4676
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-[illegible]-2

Alignment Scores:
Pred. No.:              [illegible]      Length:        [illegible]
Score:                  [illegible]      Matches:       [illegible]
Percent Similarity:     [illegible]      Conservative:  [illegible]
```

```
Best Local Similarity:  [illegible]
Query Match:            [illegible]
DB:                     [illegible]

SW [illegible] x [illegible]

Qy       [illegible]

Db       [illegible]

RESULT [illegible]
[illegible]
; Sequence [illegible], Application [illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;  APPLICANT: Li, Steve
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  TITLE OF INVENTION: NOVEL [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: [illegible]
;  CURRENT FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: [illegible]
; SEQ ID NO [illegible]
;   LENGTH: [illegible]
;   TYPE: DNA
;   ORGANISM: [illegible]
[illegible]

Alignment Scores:
Pred. No.:              [illegible]
Score:                  [illegible]
Percent Similarity:     [illegible]
Best Local Similarity:  [illegible]
Query Match:            [illegible]
DB:                     [illegible]

SW [illegible] x [illegible]

Qy       [illegible]

Db       [illegible]

RESULT [illegible]
[illegible]
; Sequence [illegible], Application [illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  TITLE OF INVENTION: NOVEL [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: [illegible]
;  CURRENT FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: [illegible]
; SEQ ID NO [illegible]
;   LENGTH: [illegible]
;   TYPE: DNA
;   ORGANISM: [illegible]
[illegible]

Alignment Scores:
Pred. No.:              [illegible]
Score:                  [illegible]
```

Command line parameters:

SUMMARIES

ALIGNMENTS

GENERAL INFORMATION:

APPLICANT:

APPLICANT:

APPLICANT:

APPLICANT:

CURRENT APPLICATION NUMBER:

CURRENT FILING DATE:

SOFTWARE:

```
[illegible]
                /gene="HI?M[illegible]"
                /note="[illegible]
                http://[illegible]
                s Sequence"
[illegible]
                /gene="HI?M[illegible]"
                /note="[illegible]
                [illegible]
                [illegible]"
                /[illegible]
                /product="hypothetical [illegible]"
                /protein_id="AAK[illegible]"
                /db_xref="[illegible]"
                /translation="M[illegible]
                AK[illegible]KMVS[illegible]
                [illegible]
                [illegible]
                [illegible]
                [illegible]
                [illegible]
                [illegible]
                [illegible]
                [illegible]
                N[illegible]NAA[illegible]"
[illegible]
                /gene="[illegible]"
                /note="[illegible]
                http://[illegible]
                s Sequence"
[illegible]
                /gene="[illegible]"
                /note="[illegible]
                [illegible]
                [illegible]
                [illegible]
                [illegible]"
                /[illegible]
                /product="[illegible]"
                /protein_id="AA[illegible]"
                /db_xref="[illegible]"
                /translation="MA[illegible]
                [illegible]
                [illegible]
                [illegible]"
[illegible]
                /gene="[illegible]"
                /note="[illegible]
                http://[illegible]
                s Sequence"
[illegible]
                /gene="[illegible]"
                /note="[illegible]

[illegible]
[illegible]      Matches [illegible]

[illegible]
```

[illegible]

[illegible] Caenorhabditis elegans [illegible]

[illegible]

Genome Sequencing Center
Department of Genetics, Washington University
[illegible]

[illegible]

[illegible] above are the result of integration and [illegible] with data [illegible] computer analysis [illegible] personal communication [illegible] of Yuji Kohara (http://[illegible]/elegans/html/[illegible]INDEX.html) and [illegible] elegans [illegible] project (http://worfdb.dfci.harvard.edu/). [illegible] proteins from BlastX analyses (http://[illegible]/), sequence conservation with [illegible] WABA alignment program ([illegible]), individual C. elegans [illegible] and personal communications with C. elegans researchers. tRNAs are predicted using the program tRNAscan-SE (Lowe, T.M. and [illegible] Nucl. Acids Res. [illegible]).

```
FEATURES
     source          [illegible]
                     /organism="Caenorhabditis elegans"
                     /strain="Bristol N2"
                     [illegible]
     gene            [illegible]
                     /gene="H27M09.1"
                     /note="for a graphical representation of this gene see
                     [illegible]"
     CDS             [illegible]
                     /gene="H27M09.1"
                     /note="contains similarity to [illegible]"
                     [illegible]
                     /codon_start=1
                     /product="Hypothetical protein H27M09.1"
                     /protein_id="AA[illegible]"
                     /db_xref="[illegible]"
                     /translation="M[illegible]HARFIATAA[illegible]VS[illegible]AVVSS[illegible]
                     [illegible]
                     [illegible]GINNGYGPVVNAEPEPQ[illegible]QQGKAG[illegible]
                     [illegible]
                     [illegible]KK[illegible]VNVN[illegible]
                     [illegible]"
     gene            [illegible]
                     /gene="H27M09.2"
                     /note="for a graphical representation of this gene see
                     [illegible]"
     CDS             [illegible]
                     /gene="H27M09.2"
                     /codon_start=1
                     /product="Hypothetical protein H27M09.2"
                     /protein_id="AA[illegible]"
                     /db_xref="[illegible]"
                     /translation="MSSS[illegible]MAENVNY[illegible]MVIA[illegible]
                     [illegible]KKR[illegible]KKSKPEKK[illegible]NYKE[illegible]NDHNNGSDEKK[illegible]
                     [illegible]KKNNVSAHKMEKSPSEE[illegible]
                     [illegible]YEEPKEEEENSHKSAESMEVEKSNSYENYKKY[illegible]
                     [illegible]YSIAVKESNSNEKSYYS[illegible]SKEKE[illegible]
                     [illegible]SKKSAKS[illegible]SKESNSV[illegible]
                     [illegible]"
```

repeat_region [illegible]

[illegible]

KEYWORDS HTG; [illegible] HTGS_DRAFT.

REFERENCE [illegible]
AUTHORS [illegible] and Waterston,R.H.

REFERENCE [illegible]
AUTHORS [illegible] and Waterston,R.H.
JOURNAL [illegible] Genome Sequencing Center [illegible] USA

COMMENT [illegible]
[illegible]
[illegible]
Project Information
[illegible]

Summary Statistics
[illegible]
[illegible] bases at least q40
[illegible] bases at least q30
[illegible] bases at least q20
[illegible]

* NOTE: [illegible] 'working draft' sequence. It currently
* [illegible] contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* [illegible] the exact sizes of the gaps are unknown.
* [illegible] be updated with the finished sequence
* [illegible] available and the accession number will
* [illegible]
* [illegible] contig of [illegible] bp in length
* [illegible] gap of unknown length
* [illegible] contig of [illegible] bp in length
* [illegible] gap of unknown length
* [illegible] contig of [illegible] bp in length
* [illegible] gap of unknown length
* [illegible] contig of [illegible] bp in length
* [illegible] gap of unknown length
* [illegible] contig of [illegible] bp in length
* [illegible] gap of unknown length
* [illegible] contig of [illegible] bp in length
* [illegible] gap of unknown length
* [illegible] contig of [illegible] bp in length
* [illegible] gap of unknown length
* [illegible] contig of [illegible] bp in length
* [illegible] gap of unknown length
* [illegible] contig of [illegible] bp in length
* [illegible] gap of unknown length
* [illegible] contig of [illegible] bp in length

[illegible]

LOCUS AC105129 [illegible] bp DNA linear PRI [illegible]
DEFINITION Homo sapiens chromosome 15, [illegible] complete sequence.
ACCESSION [illegible]
VERSION AC105129.4 GI:20330980
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to [illegible])
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 15, [illegible]
JOURNAL Unpublished
REFERENCE 2 (bases 1 to [illegible])
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., [illegible]
Anderson,S., Baldwin,J., Bastien,V., [illegible]
Brown,A., Camarata,J., Campopiano,A., Chang,J., Charlebois,R.,
[illegible]
Mahoney,M., McEwan,P., McKernan,K., Meldrim,J., [illegible]
Mihova,T., Mlenga,V., Murphy,L., Naylor,J., Nguyen,T., [illegible]
[illegible]
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
[illegible]
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted ([illegible]) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to [illegible])
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., [illegible]
Anderson,S., Baldwin,J., Bastien,V., [illegible]
Bookstaller,B., Brown,A., Camarata,J., Campopiano,A., [illegible]
[illegible]
Kumar,A., Karatas,A., [illegible]
[illegible]
Mihova,T., Mlenga,V., Murphy,L., Naylor,J., Nguyen,T., [illegible]
[illegible]
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
[illegible]
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted ([illegible]) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 4 (bases 1 to [illegible])
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., [illegible]

[illegible]

JOURNAL [illegible]
COMMENT [illegible]

FEATURES [illegible]
source [illegible]

```
            [illegible] and [illegible]
TITLE       Direct Submission
JOURNAL     Unpublished
REFERENCE   3  (bases 1 to 1[illegible])
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   4  (bases 1 to 1[illegible])
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted ([illegible]) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77[illegible], USA
COMMENT     On [illegible] 2002 this sequence version replaced [illegible]
                 --- Genome Center:
               Center: Baylor College of Medicine
               Center code: BCM
               Web site: http://www.[illegible]
               Contact: [illegible]
              --- Project Information:
               Center project name: [illegible]BA
               Center clone name: [illegible]
              --- Summary Statistics:
               Sequencing vector: Plasmid;
               Chemistry: Dye-terminator [illegible]
               Assembly program: Phrap; [illegible]
               Consensus quality: [illegible] bases at least Q4
               Consensus quality: [illegible] bases at least Q[illegible]
               Consensus quality: [illegible] bases at least Q[illegible]
               ------
            * NOTE: [illegible]
            *   (see http://[illegible])
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of [illegible] contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        [illegible]  [illegible]: contig of [illegible] bp in length
            *     1[illegible]4  [illegible]: gap of unknown length
            *     [illegible]4  [illegible]: contig of [illegible] bp in length
            *     [illegible]  [illegible]: gap of unknown length
            [illegible]
```

[illegible]

DEFINITION [illegible]

ACCESSION [illegible]

VERSION [illegible]

KEYWORDS [illegible]

SOURCE [illegible]

ORGANISM [illegible]

REFERENCE [illegible]

AUTHORS [illegible]

[illegible]

LOCUS [illegible] 116580 bp DNA [illegible]
DEFINITION Mus musculus Strain [illegible] BAC [illegible] sequence.
ACCESSION [illegible]
VERSION [illegible]
KEYWORDS HTG.
SOURCE [illegible]
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible] Mammalia; Eutheria; Rodentia; [illegible] Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 116580)
AUTHORS [illegible]
TITLE [illegible] Mouse Sequencing
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 116580)
AUTHORS [illegible]
TITLE Direct Submission
JOURNAL Submitted ([illegible]-FEB-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, [illegible] Landsdowne St, Cambridge, MA [illegible] USA
REFERENCE 3 [illegible]
AUTHORS [illegible]
TITLE Direct Submission
JOURNAL Submitted ([illegible]-MAY-[illegible]) [illegible] USA
REFERENCE 4 [illegible]
AUTHORS [illegible]
TITLE Direct Submission
JOURNAL Submitted ([illegible]) Harvard Partners [illegible] Genomics, Harvard Medical School, [illegible] Landsdowne St, Cambridge, MA [illegible] USA
REFERENCE 5 [illegible]
AUTHORS [illegible]
TITLE Direct Submission
JOURNAL Submitted ([illegible]) [illegible] USA
COMMENT [illegible]

[illegible]

FEATURES [illegible]

/translation="MAAAM...

ORGANISM Homo sapiens

REFERENCE

AUTHORS Meyers,R.A. and Williamson,M.

Millennium Pharmaceuticals, Inc. (US)

/organism="Homo sapiens"

/translation="MAAAM...

ORGANISM

REFERENCE

```
AX [illegible]
[illegible]
DEFINITION  Sequence [illegible] Patent W[illegible]
ACCESSION   AX[illegible]
VERSION     AX[illegible]
KEYWORDS    .
SOURCE      [illegible]
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
            Mammalia; Eutheria; Primates; [illegible]
REFERENCE   1  (bases 1 to [illegible])
  AUTHORS   [illegible]
  TITLE     [illegible]
  JOURNAL   Patent: W[illegible]
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
                     /organism="Homo sapiens"
                     /note="Incyte ID No: [illegible]"
BASE COUNT  [illegible]
ORIGIN

Alignment Scores:
[illegible]
Best Local Similarity:  [illegible]
Query Match:            [illegible]

SW [illegible] x AX [illegible]

[illegible]

[illegible]

DEFINITION  Homo sapiens dipeptidyl peptidase [illegible], mRNA [illegible]
            alternatively spliced.
ACCESSION   A[illegible]
VERSION     A[illegible]
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
            Mammalia; Eutheria; Primates; [illegible]
REFERENCE   1  (bases 1 to [illegible])
  AUTHORS   Abbott,C.A., Yu,D.M., Woollatt,E., Sutherland,G.R., McCaughan,G.W.
            and Gorrell,M.D.
  TITLE     Cloning, expression and [illegible]
            dipeptidyl peptidase (DPP) IV [illegible]
  JOURNAL   Eur. J. Biochem. 267 (20), [illegible] (2000)
  MEDLINE   [illegible]
REFERENCE   2  (bases 1 to [illegible])
  AUTHORS   Abbott,C.A., Yu,D., McCaughan,G.W. and Gorrell,M.D.
  TITLE     Direct Submission
  JOURNAL   Submitted [illegible]
            [illegible]
            [illegible] NSW [illegible] Australia
FEATURES             Location/Qualifiers
                     /organism="Homo sapiens"
                     /chromosome="[illegible]"
                     /tissue_type="[illegible]"
     gene            [illegible]
```

[illegible]

SUMMARY

ALIGNMENTS

```
[illegible]
KW   Hypothetical protein.
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible]

[illegible]

[illegible]
[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Stover C.K., Pham X.Q., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K., [illegible] Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT   opportunistic pathogen.";
RL   Nature [illegible].
CC   -!- SIMILARITY: BELONGS TO THE [illegible] FAMILY.
CC   [illegible]
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
[illegible]
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible]

[illegible]
```

```
RN   [illegible]
RP   SEQUENCE FROM N.A.
RX   MEDLINE [illegible] PubMed [illegible]
RT   [illegible]
RL   Proc. Natl. Acad. Sci. U.S.A. [illegible]
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; [illegible]
DR   PIR; [illegible]
KW   Transposable element; Hypothetical protein.
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches [illegible] Mismatches [illegible] Indels [illegible]

RESULT [illegible]
[illegible] STANDARD; [illegible] AA.
AC   [illegible]
DT   [illegible]
DE   [illegible]
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE [illegible] PubMed [illegible]
RA   [illegible]
RT   "Nucleotide sequence of [illegible]";
RL   Nucleic Acids Res. [illegible]
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE [illegible] PubMed [illegible]
RA   Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA   Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
RA   Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA   Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T.,
RA   Saito N., Sampei G., Seki Y., Sivasundaram S., Tagami H.,
RA   Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.;
RT   "A [illegible] DNA sequence of the Escherichia coli K-12 genome
RT   [illegible]";
RL   DNA Res. [illegible]
CC   [illegible]
CC   This SWISS-PROT entry is copyright. [illegible]
CC   between the Swiss Institute of Bioinformatics and the EMBL [illegible]

RESULT [illegible]
YAW [illegible]
ID   YAW [illegible] STANDARD; [illegible] AA.
AC   [illegible]
RP   SEQUENCE FROM N.A.
RA   [illegible]
RT   [illegible]
RL   Nature [illegible]
CC   [illegible]
```

[illegible]

RX MEDLINE [illegible]
RA [illegible] Anderson K., [illegible] Burke M.
RA [illegible]
RA [illegible] Fraser A.
RA [illegible]
RA [illegible] Kirsten [illegible]
RA [illegible]
RA [illegible] Kerkhoven [illegible] Rogers A., Saunders D., Skwarek [illegible]
RA Sims M., [illegible] Smith A., Smith M., Sonnhammer E., Staden R.
RA Sulston J., [illegible] Thomas K., Vaudin M., Vaughan K.
RA Waterston R., [illegible] Weinstock [illegible] Wilkinson-Sproat J.
RA Wohldman [illegible]
RT "[illegible] sequence from the nematode [illegible]
RT [illegible]."
RL Nature [illegible]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC [illegible]
DR EMBL [illegible]
DR WormPep [illegible]
KW Hypothetical protein.
SQ SEQUENCE [illegible] MW; [illegible]

[illegible]

KW Hypothetical protein; Nucleosome core; [illegible] DNA-binding.
KW [illegible] DNA condensation; Nuclear protein.

[illegible]

```
RP   VARIANT [illegible] AND [illegible].
RX   MEDLINE [illegible]; PubMed [illegible];
RA   Beckman [illegible], Fitzpatrick M., [illegible]
RA   [illegible]
RA   [illegible] R.W.;
RT   "[illegible] HFE [illegible]
RT   America.";
RL   Hepatology [illegible](1999).
RN   [illegible]
RP   VARIANTS HH ASP [illegible] AND [illegible].
RX   MEDLINE [illegible]; PubMed 10194428;
RA   [illegible]
RT   "HFE mutations analysis in [illegible]
RT   S65C implication in mild form of hemochromatosis.";
RL   Blood [illegible]:2502-250[illegible](1999).
RN   [14]
RP   VARIANTS HH [illegible]; AR [illegible] AND [illegible].
RX   MEDLINE [illegible]; PubMed [illegible];
RA   Barton J.C., Sawada-Hirai R., Rothenberg B.E., Acton R.T.;
RT   "Two novel missense mutations of the HFE gene ([illegible] and [illegible]) and
RT   [illegible] of the S65C mutation in Alabama [illegible]
RT   [illegible].";
RL   Blood Cells Mol. Dis. 25:14[illegible]-15[illegible](1999).
RN   [illegible]
RP   VARIANTS VE [illegible] AND [illegible]; VARIANT HH [illegible] AND VARIANTS [illegible]
RP   [illegible].
RX   MEDLINE [illegible]; PubMed [illegible];
RA   de Villiers J.N.P., Hillermann R., [illegible]
RT   "Spectrum of mutations in the HFE gene [illegible]
RT   [illegible].";
RL   Hum. Mol. Genet. [illegible]:1517-152[illegible]([illegible]).
RN   [illegible]
RP   VARIANTS HH AS[illegible] AND [illegible].
RX   MEDLINE [illegible]; PubMed [illegible];
RA   Merryweather-Clarke A.T., [illegible]
RA   Nordlund-Pedersen [illegible], Robson K.J.H.;
RT   "A [illegible]
RT   [illegible] in Scandinavian populations.";
RL   Hum. Mutat. [illegible](1999).
RN   [illegible]
RP   VARIANT HH [illegible].
RA   [illegible]
RT   "A [illegible] missense mutation [illegible] in the HFE gene with a possible [illegible]
RT   in hereditary haemochromatosis.";
RL   Hum. Mutat. [illegible](1999).
RN   [illegible]
RP   VARIANT LYS-2[illegible].
RX   MEDLINE 2[illegible]; PubMed 1[illegible];
RA   [illegible] S.[illegible];
RT   "[illegible] (E[illegible]K and V[illegible]) [illegible]
RT   [illegible].";
RL   Hum. Mutat. 1[illegible]:120-120(2000).
CC   -!- FUNCTION: BINDS TO TRANSFERRIN RECEPTOR [illegible] AND [illegible]
CC       [illegible] TRANSFERRIN.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: 4 isoforms; [illegible] (shown here), [illegible]
CC       [illegible] are produced by alternative splicing.
CC   -!- TISSUE SPECIFICITY: IN ALL TISSUES TESTED EXCEPT BRAIN.
CC   -!- DISEASE: DEFECTS IN HFE ARE A CAUSE OF HEREDITARY HEMOCHROMATOSIS
CC       (HH). HH IS AN AUTOSOMAL RECESSIVE INBORN ERROR OF IRON
CC       METABOLISM FREQUENT AMONG CAUCASIANS. HH IS CHARACTERIZED BY
CC       ABNORMAL INTESTINAL IRON ABSORPTION AND PROGRESSIVE INCREASE
CC       [illegible] BODY IRON WHICH RESULTS IN MULTIPLE [illegible]
CC       COMPLICATIONS INCLUDING CIRRHOSIS, [illegible], DIABETES,
CC       ENDOCRINE DYSFUNCTIONS, ARTHROPATHY AND [illegible]
CC       [illegible], SINCE THE DISEASE [illegible] CAN BE EFFECTIVELY
CC       PREVENTED BY REGULAR [illegible], EARLY [illegible]
CC       [illegible] A NORMAL LIFE EXPECTANCY [illegible] THE AFFECTED
CC       [illegible].
CC   -!- DISEASE: DEFECTS IN HFE ARE A CAUSE OF [illegible]
CC       [illegible] CHARACTERIZED [illegible]
CC       [illegible] AND PRESENCE OF LARGE AMOUNTS OF [illegible]
CC       [illegible] IRON OVERLOAD IS OFTEN PRESENT IN ASSOCIATION WITH VARI[illegible]
```

```
CC   [illegible]
DR   EMBL; [illegible]
DR   EMBL; [illegible]
DR   EMBL; [illegible]
DR   EMBL; [illegible]
DR   EMBL; [illegible]
DR   EMBL; [illegible]
DR   EMBL; [illegible]
DR   EMBL; [illegible]
DR   EMBL; [illegible]
DR   EMBL; [illegible]
DR   EMBL; [illegible]
[illegible]
RESULT [illegible]
[illegible]
FT   [illegible]
DE   [illegible]
RX   [illegible]
RA   [illegible]
RT   [illegible]
RL   [illegible]
CC   [illegible]
```

```
DR   EMBL; [illegible]; AA[illegible].
DR   EMBL; [illegible]; AA[illegible].
DR   EMBL; [illegible]; AAB[illegible].
DR   EMBL; [illegible]; BAA[illegible].
DR   EMBL; [illegible]; BAA[illegible].
DR   [illegible]; [illegible]; [illegible].
DR   InterPro; IPR00[illegible]; ATP-bd_ATPase.
DR   InterPro; IPR[illegible]; [illegible].
DR   InterPro; IPR[illegible]; [illegible].
DR   InterPro; IPR[illegible]; PAS_[illegible].
DR   Pfam; PF00989; PAS; 1.
DR   Pfam; PF02518; HATPase_c; 1.
DR   PRINTS; PR[illegible]; [illegible].
DR   SMART; SM[illegible]; HATPase_c; 1.
DR   SMART; SM00091; PAS; 1.
DR   PROSITE; PS[illegible]; HIS_KIN[illegible]; 1.
DR   PROSITE; PS[illegible]; PAS; FACTOR_NEG[illegible].
KW   Sensory transduction; Transferase; Kinase; [illegible].
KW   Transmembrane; Inner membrane; [illegible] protein; [illegible].
FT   [illegible] MAIN   [illegible]   [illegible]   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   [illegible]   [illegible]   POTENTIAL.
FT   [illegible] MAIN   [illegible]   [illegible]   PERIPLASMIC (POTENTIAL).
FT   TRANSMEM   [illegible]   [illegible]   POTENTIAL.
FT   [illegible] MAIN   [illegible]   [illegible]   CYTOPLASMIC (POTENTIAL).
FT   [illegible] MAIN   [illegible]   [illegible]   PAS.
FT   [illegible] MAIN   [illegible]   [illegible]   HISTIDINE KINASE.
FT   MOD_RES   347   347   PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   [illegible]   [illegible]   [illegible]   [illegible]
FT   [illegible] (IN REF. [illegible]).
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible] CRC64;

  Query Match [illegible]; Score 28; [illegible]; Length [illegible];
  Best Local Similarity [illegible]; Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy   [illegible] ERHSIR [illegible]
Db   [illegible] ERHSIR [illegible]

RESULT [illegible]
[illegible]_BACSU
ID   [illegible]_BACSU   STANDARD;   PRT;   [illegible] AA.
AC   [illegible];
DT   [illegible] (Rel. [illegible], Created)
DT   [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] (Rel. [illegible], Last annotation update)
DE   ComE operon protein [illegible].
GN   [illegible].
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Hahn J., Inamine G., Kozlov Y., Dubnau D.A.;
RT   "Characterization of comE, a late competence operon of Bacillus
RT   subtilis required for the binding and uptake of transforming DNA.";
RL   Mol. Microbiol. [illegible].
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=[illegible];
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Mizuno M., Masuda S., Takemaru K., [illegible];
RA   Kobayashi Y.;
RT   "Systematic sequencing of the [illegible] kb [illegible] region of
RT   the Bacillus subtilis genome containing the [illegible]
RT   operon [illegible].";
RL   Microbiology [illegible].
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   STRAIN=[illegible];
RX   MEDLINE=[illegible]; PubMed=[illegible];
```

[illegible]

RP SEQUENCE FROM N.A.
RA Namath A., [illegible] Roberts D., [illegible] Garcia E.W.;
RL Submitted ([illegible] JAN [illegible]) to the EMBL/GenBank/DDBJ databases.
[illegible]
RP SEQUENCE FROM N.A.
RT STRAIN K12;
[illegible]
RA Kimura S., [illegible] Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., [illegible] Nakamura Y., Nishimura [illegible] Nishi [illegible]
[illegible]
CC -!- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM [illegible]
[illegible]

```
DR   SMART; SM[illegible]
DR   SMART; SM[illegible]; SH2; [illegible]
DR   SMART; SM[illegible]; TyrKc; [illegible]
DR   PROSITE; PS5[illegible]; FCH; [illegible]
DR   PROSITE; PS[illegible]; PROTEIN_KINASE_ATP; [illegible]
DR   PROSITE; PS[illegible]; PROTEIN_KINASE_DOM; [illegible]
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS[illegible]; SH2; 1.
KW   Transferase; Tyrosine-protein kinase; [illegible]; ATP-binding;
KW   Phosphorylation.
FT   DOMAIN        [illegible]     [illegible]       FCH
FT   DOMAIN       45[illegible]    5[illegible]7      SH2
FT   DOMAIN       5[illegible]     [illegible]       PROTEIN KINASE.
FT   NP_BIND      5[illegible]     [illegible]       ATP (BY SIMILARITY).
FT   BINDING      588    588       ATP (BY SIMILARITY).
FT   ACT_SITE     [illegible]     [illegible]       BY SIMILARITY.
FT   MOD_RES      7[illegible]     7[illegible]       PHOSPHORYLATION (A[illegible]) (BY SIMILARITY).
FT   CONFLICT     747    747       [illegible] (IN REF. [illegible]).
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match              [illegible];  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy        [illegible]
              [illegible]
Db        [illegible]


RESULT 2
RPO1_ARATH
ID   RPO1_ARATH     STANDARD;      PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] (Rel. [illegible], Created)
DT   [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] (Rel. [illegible], Last annotation update)
DE   DNA-directed RNA polymerase, chloroplast precursor (EC 2.7.7.6).
GN   RPO[illegible] OR AT2G[illegible] OR F27[illegible].
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; [illegible]
OC   [illegible]; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible], [illegible] W.[illegible] A.;
RT   "Mitochondrial and chloroplast phage-type RNA polymerases in
RT   Arabidopsis."
RL   Science 277:[illegible](1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Lin X., Kaul S., [illegible] S., Shea T., [illegible] M., [illegible]
RA   [illegible] C.Y., Marn [illegible] M., [illegible] C.[illegible], [illegible] M.E., [illegible] T.V
RA   [illegible] K.A., [illegible] [illegible]
RA   [illegible] K.S., [illegible] L.A., [illegible] M., [illegible] Van Aken S., [illegible]
RA   [illegible] C.F., Adams M.[illegible], [illegible] A.[illegible], [illegible]
RA   [illegible] M., [illegible]
RA   Nierman W.[illegible], White [illegible], [illegible] J.A., [illegible], Fraser C.M.,
RA   Venter [illegible];
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:[illegible](1999).
CC   -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC       OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC       SUBSTRATES.
CC   -!- CATALYTIC ACTIVITY: N [illegible] = N diphosphate +
CC       RNA(N)
CC   -!- SUBCELLULAR LOCATION: Chloroplast.
CC   -!- SIMILARITY: BELONGS TO THE PHAGE AND M[illegible] RNA POLYMERASE [illegible]
```

```
[illegible]
CC   [illegible]
DR   EMBL; [illegible]
DR   EMBL; [illegible]
DR   EMBL; [illegible]
KW   [illegible]
KW   [illegible]
FT   [illegible]
SQ   SEQUENCE [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy   [illegible]

Db   [illegible]


RESULT 3
[illegible]
ID   [illegible]
AC   [illegible]
DT   [illegible]
DT   [illegible]
DT   [illegible]
DE   [illegible]
GN   [illegible]
OS   [illegible]
OC   [illegible]
OC   [illegible]
OC   [illegible]
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   [illegible]
RC   [illegible]
RX   MEDLINE=[illegible]
RA   [illegible]
RT   [illegible]
RT   [illegible]
RL   [illegible]
RN   [2]
RP   [illegible]
RC   [illegible]
RX   MEDLINE=[illegible]
RA   [illegible]
RA   [illegible]
RT   [illegible]
RT   [illegible]
RL   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
```

[illegible]

SUMMARIES

[illegible]

ALIGNMENTS

[illegible]

RX MEDLINE [illegible]
RA WILKS A.F. [illegible]
RT "[illegible] polymerase chain reaction to cloning [illegible]"
RL [illegible]
CC [illegible] ATP [illegible] protein tyrosine [illegible] ATP [illegible]
CC [illegible] SIMILARITY: [illegible] THE TYR FAMILY OF PROTEIN KINASES [illegible]
CC [illegible]
CC -!- SIMILARITY: [illegible] SH2 DOMAIN.
CC -!- SIMILARITY: [illegible] SH3 DOMAIN.
CC [illegible]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL [illegible]
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. [illegible]
CC entities requires a license agreement (See http://[illegible]
CC [illegible]

DR EMBL; [illegible]
DR EMBL; [illegible]

[illegible]

```
;       CITY: New York
;       STATE: New York
;       COUNTRY: USA
;       ZIP: [illegible]
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: Windows 95
;       SOFTWARE: FastSEQ for Windows Version [illegible]
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/ [illegible]A
;       FILING DATE: [illegible] APR [illegible]
;       CLASSIFICATION: [illegible]14
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: US [illegible]
;       FILING DATE: [illegible] MAY [illegible]
;       CLASSIFICATION: [illegible]
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: US [illegible]
;       FILING DATE: [illegible] APR [illegible]
;       CLASSIFICATION: [illegible]
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: US [illegible]
;       FILING DATE: [illegible] APR 1996
;       CLASSIFICATION: 514
;     ATTORNEY/AGENT INFORMATION:
;       NAME: [illegible], Brian M.
;       REGISTRATION NUMBER: [illegible]
;       REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: [illegible]
;       TELEFAX: [illegible]
;       TELEX: [illegible]
;   INFORMATION FOR SEQ ID NO: [illegible]
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: [illegible] amino acids
;       TYPE: amino acid
;       TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-[illegible]

[illegible]

RESULT [illegible]
[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]
;   GENERAL INFORMATION:
;     APPLICANT: [illegible]
;     APPLICANT: [illegible], Dennis [illegible]
;     APPLICANT: [illegible], [illegible] N.
;     APPLICANT: [illegible], Andreas
;     APPLICANT: [illegible], David
;     APPLICANT: [illegible], [illegible]
;     APPLICANT: [illegible], Robert K.
;     TITLE OF INVENTION: Hereditary [illegible]
;     NUMBER OF SEQUENCES: 44
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: [illegible]
;       STREET: [illegible] Avenue of the Americas
;       CITY: New York
;       STATE: New York
;       COUNTRY: USA
;       ZIP: [illegible]
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
```

```
[illegible]
```

```
;     TELEFAX: (415) [illegible]
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 348 amino acids
;       TYPE: amino acid
;       TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible]

Qy [illegible]

Db [illegible]


RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;    APPLICANT: Thomas, [illegible]
;    APPLICANT: Drayna, Dennis [illegible]
;    APPLICANT: Feder, John N.
;    APPLICANT: [illegible], Andreas
;    APPLICANT: [illegible]
;    APPLICANT: Tsuchihashi, [illegible]
;    APPLICANT: Wolff, Roger K.
;    TITLE OF INVENTION: Hereditary [illegible]
;    NUMBER OF SEQUENCES: 44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Townsend and Townsend and Crew LLP
;      STREET: [illegible]
;      CITY: San Francisco
;      STATE: California
;      COUNTRY: USA
;      ZIP: [illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.[illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/[illegible]
;      FILING DATE: [illegible] MAY [illegible]
;      CLASSIFICATION: [illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Smith, William M.
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (415) [illegible]
;      TELEFAX: (415) [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 348 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;    MOLECULE TYPE: protein
US-[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible]

Qy [illegible]

Db [illegible]
```

[illegible]

```
;       OPERATING SYSTEM: Windows
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: [illegible]
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: [illegible]
;     ATTORNEY/AGENT INFORMATION:
;       NAME: [illegible] Brian M
;       REGISTRATION NUMBER: 28,462
;       REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: [illegible]
;       TELEFAX: [illegible]
;       TELEX: [illegible] PENNIE
;   INFORMATION FOR SEQ ID NO: [illegible]
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: [illegible]
;       TYPE: amino acid
;       STRANDEDNESS: [illegible]
;       TOPOLOGY: [illegible]
;     MOLECULE TYPE: peptide
[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

[illegible]

[illegible]


[illegible]
[illegible]
; Sequence [illegible] Application [illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;     APPLICANT: [illegible]
;     APPLICANT: [illegible]
;     APPLICANT: [illegible]
;     TITLE OF INVENTION: METHODS AND COMPOSITIONS [illegible]
;     TITLE OF INVENTION: DIAGNOSIS AND TREATMENT [illegible]
;     TITLE OF INVENTION: AND [illegible] DEFICIENCY [illegible]
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Pennie & Edmonds [illegible]
;       STREET: [illegible] Avenue of the Americas
;       CITY: New York
;       STATE: NY
;       COUNTRY: USA
;       ZIP: [illegible]
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Diskette
;       COMPUTER: IBM Compatible
;       OPERATING SYSTEM: Windows
;       SOFTWARE: [illegible] for Windows Version [illegible]
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: [illegible]
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: [illegible]
;     ATTORNEY/AGENT INFORMATION:
;       NAME: [illegible] Brian M
;       REGISTRATION NUMBER: 28,462
;       REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: [illegible]
```

[illegible]

Maximum Match
Minimum Match
Listing first 45 summaries

; [illegible] INFORMATION: Xaa [illegible]
; NAME/KEY: [illegible]
; LOCATION: (1)..(?)
; OTHER INFORMATION: Xaa [illegible] any of the [illegible]
[illegible]

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

[illegible]

RESULT [illegible]
US-[illegible]-A
; Sequence [illegible], Application US/[illegible]A
; Patent No. [illegible]A
; GENERAL INFORMATION:
; APPLICANT: [illegible]
; APPLICANT: [illegible], Paige W.
; APPLICANT: [illegible], Peter W.
; TITLE OF INVENTION: MUTATION [illegible] BIOSYNTHESIS [illegible]
; FILE REFERENCE: [illegible]
; CURRENT APPLICATION NUMBER: US/[illegible]A
; CURRENT FILING DATE: [illegible]
; PRIOR APPLICATION NUMBER: US [illegible]
; PRIOR FILING DATE: [illegible]
; NUMBER OF SEQ ID NOS: [illegible]
; SOFTWARE: PatentIn [illegible]
; SEQ ID NO [illegible]
; LENGTH: [illegible]
; TYPE: [illegible]
; ORGANISM: Streptococcus mutans
US-[illegible]A-[illegible]

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

[illegible]

RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]A
; GENERAL INFORMATION:
; APPLICANT: [illegible], Robert
; APPLICANT: [illegible], Karl [illegible]
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: [illegible]
; APPLICANT: [illegible]
; APPLICANT: Yamamoto, Robert [illegible]
; APPLICANT: X[illegible], Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: [illegible]
; CURRENT APPLICATION NUMBER: US/[illegible]
; CURRENT FILING DATE: [illegible]
; PRIOR APPLICATION NUMBER: [illegible]
; PRIOR FILING DATE: [illegible]
; PRIOR APPLICATION NUMBER: [illegible]/[illegible],818
; PRIOR FILING DATE: [illegible]
; PRIOR APPLICATION NUMBER: [illegible]
; PRIOR FILING DATE: [illegible]
; PRIOR APPLICATION NUMBER: [illegible]
; PRIOR FILING DATE: [illegible]
; PRIOR APPLICATION NUMBER: [illegible]/[illegible],625

[illegible]

; FILE REFERENCE: [illegible]
; CURRENT APPLICATION NUMBER: US/[illegible]
; CURRENT FILING DATE: [illegible]
; PRIOR APPLICATION NUMBER: [illegible]
; PRIOR FILING DATE: [illegible]
; NUMBER OF SEQ ID NOS: [illegible]
; SOFTWARE: FastSEQ for Windows Version [illegible]
; SEQ ID NO [illegible]
; LENGTH: [illegible]
; TYPE: [illegible]
; ORGANISM: Homo sapiens
US-[illegible]

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

[illegible]

RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
; APPLICANT: Li, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Rozieres, Pierre
; APPLICANT: Jarlier, Jean Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: [illegible]
; CURRENT APPLICATION NUMBER: US/[illegible]
; CURRENT FILING DATE: 20[illegible]
; PRIOR APPLICATION NUMBER: US [illegible]
; PRIOR FILING DATE: [illegible]
; NUMBER OF SEQ ID NOS: [illegible]
; SOFTWARE: FastSEQ for Windows Version [illegible]
; SEQ ID NO [illegible]
; LENGTH: [illegible]
; TYPE: PRT
; ORGANISM: Homo sapiens
US-[illegible]

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

[illegible]

RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
; APPLICANT: Li, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Rozieres, Pierre
; APPLICANT: Jarlier, Jean Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: [illegible]
; CURRENT APPLICATION NUMBER: US/[illegible]
; CURRENT FILING DATE: 20[illegible]
; PRIOR APPLICATION NUMBER: US [illegible]
; PRIOR FILING DATE: [illegible]
; NUMBER OF SEQ ID NOS: [illegible]
; SOFTWARE: FastSEQ for Windows Version [illegible]
; SEQ ID NO [illegible]
; LENGTH: [illegible]

; [illegible]
; ORGANISM: [illegible]
US-[illegible]

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible]

[illegible]

RESULT [illegible]
[illegible]
; Sequence [illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
; APPLICANT: [illegible]
; TITLE OF INVENTION: [illegible]
; FILE REFERENCE: [illegible]
; CURRENT APPLICATION NUMBER: [illegible]
; PRIOR APPLICATION NUMBER: [illegible]
; PRIOR FILING DATE: [illegible]
; NUMBER OF SEQ ID NOS: [illegible]
; SOFTWARE: [illegible]
; SEQ ID NO [illegible]
; LENGTH: [illegible]
; TYPE: [illegible]
; ORGANISM: [illegible]

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible]

[illegible]

RESULT [illegible]
[illegible]
; Sequence [illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
; APPLICANT: [illegible]
; TITLE OF INVENTION: [illegible]
; FILE REFERENCE: [illegible]
; CURRENT APPLICATION NUMBER: [illegible]
; PRIOR FILING DATE: [illegible]
; NUMBER OF SEQ ID NOS: [illegible]
; SOFTWARE: [illegible]
; SEQ ID NO [illegible]
; LENGTH: [illegible]
; TYPE: [illegible]
; ORGANISM: [illegible]

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible]

[illegible]

SUMMARIES

ALIGNMENTS

```
[illegible] AAB[illegible].
XX
[illegible] New polypeptides [illegible], member of human prolyl [illegible]peptidase family,
[illegible] useful as [illegible] targets and therapeutic agents [illegible]
[illegible] [illegible] and [illegible]
XX
PS [illegible]; Page [illegible]; [illegible]pp; English.
XX
[illegible] This invention relates to [illegible] human prolyl
[illegible] [illegible]peptidase. Which is cytostatic, antidiabetic, anti[illegible],
[illegible] neuroprotective, antithyroid, dermatological, antipsoriatic,
[illegible] antiasthmatic, ophthalmological, antiinflammatory, nootropic,
[illegible] antiparkinsonian, anticonvulsant, [illegible],
[illegible] [illegible] and
[illegible] metabolic in its action. Uses include gene therapy, expression [illegible]
[illegible] activity of [illegible] protein modulator, [illegible]
[illegible] [illegible] and can be used in preventing, treating
[illegible] [illegible] proliferative [illegible] differentiative disorders
[illegible] The [illegible] molecules can act as novel diagnostic targets and therapeutic
[illegible] agents for controlling disorders associated with [illegible]
[illegible] [illegible] of peptide [illegible]
[illegible] [illegible] and proliferative [illegible]
[illegible] reproductive, neurological and cardiovascular functions. The [illegible]
[illegible] molecules are thus useful for treating and preventing cellular
[illegible] proliferative and differentiative disorders [illegible]
[illegible] [illegible] disorders such as autoimmune disease [illegible]
[illegible] mellitus, arthritis, multiple sclerosis, asthma, [illegible] disease
[illegible] [illegible] disorders, demyelinating diseases, [illegible] disorders and
[illegible] [illegible] pain disorders. This sequence represents the [illegible]
[illegible] [illegible] human prolyl [illegible]peptidase.
XX
SQ Sequence [illegible] AA;

 [illegible] Match [illegible] Score [illegible]
 Best Local Similarity [illegible] Pred. No. [illegible]
 Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

[illegible] [illegible]VL[illegible]
 [illegible]
Db [illegible] [illegible]VL[illegible] [illegible]


RESULT [illegible]
AAB[illegible]
ID AAB[illegible] standard; peptide; [illegible] AA.
XX
AC AAB[illegible];
XX
DT [illegible]-SEP-[illegible] (first entry)
XX
DE Human dipeptidyl peptidase [illegible] (DP[illegible]) peptide #[illegible].
XX
KW Human; dipeptidyl peptidase [illegible]; [illegible] diabetes; [illegible]
KW [illegible]
KW [illegible] antidiabetic; antiinflammatory; [illegible] suppressive;
KW antiviral; enzyme.
XX
OS [illegible] sapiens.
XX
PN WO[illegible]-A1.
XX
PD [illegible]-MAY-[illegible].
XX
PF [illegible]; [illegible]WO-AU01[illegible].
XX
PR [illegible] AU-[illegible].
XX
PA (UNSY-) UNIV SYDNEY.
XX
PI [illegible]
XX
DR WPI; [illegible]/[illegible].
XX
```

```
[illegible]
[illegible]
[illegible]
[illegible]
XX
PS [illegible]
XX
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
XX
SQ Sequence [illegible] AA;

 [illegible] Match [illegible]
 Best [illegible]
 Matches [illegible]

[illegible]
 [illegible]
Db [illegible] [illegible]VL[illegible]

Search [illegible]
[illegible]
```

XX

[illegible]

XX

The present invention relates to the isolation of novel human [illegible] proteins referred to as dipeptidyl peptidase IV [illegible] related proteins [illegible] and nucleic acids encoding them are useful for treating infections such as fungal, bacterial, protozoan and viral infections, particularly infections caused by human immunodeficiency virus (HIV-1 or HIV-2), [illegible] diabetes, [illegible] Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, [illegible] psychotic and neurological disorders (e.g. anxiety, [illegible] schizophrenia), and dyskinesias. These may also be used in diagnostic [illegible] for the treatment of [illegible]

XX

Sequence [illegible] aa;

[illegible]

XX

KW [illegible]

KW [illegible]

KW [illegible]

KW [illegible]

KW [illegible]

KW [illegible]

XX

[illegible]

XX

Claim [illegible]

XX

The present [illegible] relates to [illegible] new human proteins referred to [illegible] The [illegible] polynucleotides and polypeptides of the invention are useful in the diagnosis, treatment and prevention of [illegible] myocardial infarction, autoimmune/inflammatory (e.g. acquired immunodeficiency syndrome (AIDS), allergies and osteoarthritis), proliferative [illegible] developmental [illegible] muscular [illegible] epithelial (e.g. dermatitis), neurological (e.g. epilepsy and Alzheimer's disease) and reproductive (e.g. infertility and endometriosis) [illegible] Numerous other examples [illegible] The present protein sequence represents the human protein [illegible] protein of the invention.

XX

Sequence [illegible]

Query Match [illegible] Score [illegible] Length [illegible]

[illegible]

XX

KW [illegible]

KW [illegible]

KW [illegible]

KW [illegible]

KW [illegible]

KW [illegible]

XX

[illegible]

```
XX
OS   Homo sapiens.
XX
FH   Key                 Location/Qualifiers
FT   Active-site         [illegible]
FT                       /note= "Forms part of Ser Asp [illegible] catalytic triad"
FT   Active-site         [illegible]
FT                       /note= "Forms part of Ser Asp [illegible] catalytic triad"
FT   Active-site         [illegible]
FT                       /note= "Forms part of Ser Asp His catalytic triad"
XX
PN   W[illegible]AU.
XX
PD   [illegible]-MAR-[illegible].
XX
PF   [illegible]-SEP-[illegible]   [illegible]W-AU[illegible].
XX
PR   [illegible]-SEP-[illegible]   [illegible]AU[illegible].
PR   [illegible]-FEB-[illegible]   [illegible]AU[illegible].
XX
PA   (UNSY ) UNIV SYDNEY.
XX
PI   [illegible] MD;
XX
DR   WPI; [illegible]/[illegible].
DR   N-PSDB; AA[illegible].
XX
PT   New human dipeptidyl amino peptidase (DPP8), useful for [illegible]
PT   substrates, identifying inhibitors of DPP8 catalytic activity which
PT   have therapeutic uses, and for detecting activated T cells
XX
PS   Claim [illegible]; Fig [illegible]; [illegible]pp; English.
XX
CC   This sequence represents human dipeptidyl amino peptidase (DPP8).
CC   DPP8 has substrate specificity for H-Ala-Pro-pNA [illegible] pNA and
CC   H-Arg-Pro-pNA. Therefore, it is a prolyl [illegible]peptidase and a
CC   dipeptidyl peptidase, because it is capable of hydrolysing the
CC   peptide bond C-terminal to Pro in each of these compounds. DPP8
CC   is homologous with human DPPIV. DPP8 is useful for cleaving a
CC   substrate, and for detecting an activated T cell, which involves
CC   measuring the level of DPP8 gene expression in a T cell. The level
CC   of DPP8 expression is detected by detecting the amount of DPP8 [illegible]
CC   in the cell. It is also used for identifying a [illegible]
CC   of inhibiting the cleavage of the substrate [illegible] DPP8. [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible] dependent diabetes [illegible]
CC   [illegible]
CC   as immunosuppressants.
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible]

Qy        [illegible]
Db     [illegible]

RESULT [illegible]
AA[illegible]
ID   AA[illegible] standard; Protein; [illegible] AA.
XX
AC   AA[illegible];
XX
DT   [illegible]-FEB-[illegible]  (first entry)
XX
DE   Human dipeptidyl peptidase 8 (DPP8) protein.
XX
KW   Human; dipeptidyl peptidase; [illegible]
KW   [illegible]
```

[illegible]

```
CC   infections caused by human immunodeficiency virus (HIV-1 or HIV 2),
CC   pain, diabetes, precocious puberty, infertility, obesity, anorexia,
CC   bulimia, Parkinson's disease, acute heart failure, hypotension,
CC   hypertension, urinary retention, osteoporosis, angina pectoris,
CC   stroke, ulcers, asthma, allergies, cancers, migraine, vomiting,
CC   psychotic and neurological disorders (e.g. anxiety, dementia, or
CC   schizophrenia) and dyskinesias. These may also be used in discovering
CC   therapeutic agents for the treatment of reproductive, inflammatory and
CC   metabolic disorders. ABG61591-ABG61612 represent human DPRP proteases.
XX
SQ   Sequence   [illegible] AA;

  Query Match             100.0%;  Score [illegible];  DB [illegible];  Length 6[illegible];
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps

Qy          1 EVIQEEF 7
              |||||||
Db        255 EVIQEEF 261


RESULT 6
ABG615[illegible]6
ID   ABG615[illegible]6 standard; Protein; 66[illegible] AA.
XX
AC   ABG615[illegible]6;
XX
DT   12 AUG 2002 (first entry)
XX
DE   Human DPRP [illegible] splice variant #3.
XX
KW   Human; serine protease; dipeptidyl peptidase IV related protein; DPRP;
KW   DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW   diabetes; infertility; obesity; anorexia; Parkinson's disease; [illegible]
KW   heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW   ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW   dyskinesia; reproductive disorder; inflammatory disorder;
KW   metabolic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200231134-A2.
XX
PD   18-APR-2002.
XX
PF   12-OCT-2001; 2001WO-US31871.
XX
PR   12-OCT-2000; 2000US-240117P.
XX
PA   (FERR ) FERRING BV.
XX
PI   Qi S,  Akinsanya KO,  Riviere PJ,  Junien J;
XX
DR   WPI; 2002-444178/47.
DR   N-PSDB; ABK83327.
XX
PT   New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT   the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT   viral infections, cancers, allergies, neurological disorders, or pain
PT   -
XX
PS   Disclosure; Page 63-65; 118pp; English.
XX
CC   The present invention relates to the isolation of novel human serine
CC   proteases referred to as dipeptidyl peptidase IV (DPPIV)-related
CC   proteins (DPRP). The dipeptidyl peptidase IV related proteins (DPRP),
CC   and nucleic acids encoding them are useful for treating infections
CC   such as fungal, bacterial, protozoan and viral infections, particularly
CC   infections caused by human immunodeficiency virus (HIV-1 or HIV 2),
CC   pain, diabetes, precocious puberty, infertility, obesity, anorexia,
CC   bulimia, Parkinson's disease, acute heart failure, hypotension,
CC   hypertension, urinary retention, osteoporosis, angina pectoris,
CC   stroke, ulcers, asthma, allergies, cancers, migraine, vomiting,
```

```
CC   psychotic and neurological disorders [illegible]
CC   schizophrenia) and dyskinesias. [illegible]
CC   therapeutic agents [illegible]
CC   metabolic disorders. [illegible] represent human [illegible]
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy          1 EVIQEEF [illegible]
              |||||||
Db          [illegible] EVIQEEF [illegible]


RESULT 7
ABG6[illegible]
ID   ABG6[illegible] standard; [illegible]
XX
AC   ABG6[illegible]
XX
DT   12 AUG 2002 (first entry)
XX
DE   Human DPRP [illegible] splice variant #[illegible]
XX
KW   Human; serine protease; [illegible]
KW   DPPIV; infection; human [illegible] HIV-1; HIV-2; [illegible]
KW   diabetes; infertility; [illegible]
KW   heart failure; hypertension; [illegible]
KW   ulcer; allergy; cancer; [illegible]
KW   dyskinesia; reproductive [illegible] inflammatory disorder;
KW   metabolic disorder.
XX
OS   Homo sapiens.
XX
PN   WO2002[illegible] A2.
XX
PD   18-APR-2002.
XX
PF   [illegible]
XX
PR   [illegible]
XX
PA   (FERR ) FERRING BV.
XX
PI   Qi S,  Akinsanya K[illegible]
XX
DR   WPI; 2002-444178/47.
DR   N-PSDB; ABK8[illegible]
XX
PT   New dipeptidyl peptidase [illegible]
PT   the proteins, useful [illegible]
PT   viral infections, [illegible]
PT   -
XX
PS   Disclosure; Page [illegible]
XX
CC   The present invention [illegible]
CC   proteases referred to [illegible] dipeptidyl peptidase IV (DPPIV)[illegible]
CC   proteins (DPRP). [illegible]
CC   and nucleic acids [illegible]
CC   such as fungal, [illegible]
CC   infections caused by [illegible] (HIV-[illegible]
CC   pain, diabetes, [illegible]
CC   bulimia, Parkinson's [illegible]
CC   hypertension, urinary [illegible]
CC   stroke, ulcers, asthma, [illegible]
CC   psychotic and [illegible]
CC   schizophrenia) and [illegible]
CC   therapeutic agents [illegible]
CC   metabolic disorders. [illegible] represent [illegible]
XX
```

[illegible]

XX

[illegible]

XX

KW [illegible] dipeptidyl peptidase IV related protein [illegible]
KW [illegible] human immunodeficiency virus; HIV [illegible] HIV [illegible]
KW [illegible] Parkinson's disease [illegible]
KW heart [illegible] urinary [illegible] osteoporosis [illegible]
KW [illegible] disorder [illegible]
KW [illegible] inflammatory disorder [illegible]
KW [illegible]

XX

[illegible]

XX

[illegible]

XX

PI [illegible]

XX

PT New dipeptidyl peptidase IV related proteins and nucleic acids [illegible]
PT [illegible] treating e.g. fungal, bacterial, protozoal and
PT viral infections, cancers, allergies, neurological disorders [illegible]

XX

PS [illegible] English.

XX

CC The present invention relates to the isolation of novel human [illegible]
CC proteins [illegible] dipeptidyl peptidase IV (DPP IV) related
CC proteins [illegible] dipeptidyl peptidase IV related proteins [illegible]
CC and nucleic acids encoding them are useful for treating infections
CC [illegible] fungal [illegible] protozoal and viral infections [illegible]

```
DR   N-PSDB; ABK83328.
XX
PT   New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT   the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT   viral infections, cancers, allergies, neurological disorders, or [illegible]
PT   [illegible]
XX
PS   Disclosure; Page 66-67; 113pp; English.
XX
CC   The present invention relates to the isolation of novel human serine
CC   proteases referred to as dipeptidyl peptidase IV (DPPIV) related
CC   proteins (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP)
CC   and nucleic acids encoding them are useful for treating infections
CC   such as fungal, bacterial, protozoan and viral infections, particularly
CC   infections caused by human immunodeficiency virus (HIV 1 or HIV 2),
CC   pain, diabetes, precocious puberty, infertility, obesity, anorexia,
CC   bulimia, Parkinson's disease, acute heart failure, hypotension,
CC   hypertension, urinary retention, osteoporosis, angina pectoris,
CC   stroke, ulcers, asthma, allergies, cancers, migraine, vomiting,
CC   psychotic and neurological disorders (e.g. anxiety, dementia, or
CC   schizophrenia), and dyskinesias. These may also be used in discovering
CC   therapeutic agents for the treatment of reproductive, inflammatory and
CC   metabolic disorders. ABG61591-ABG61612 represent human DPRP proteins.
XX
SQ   Sequence   858 AA;

  Query Match             100.0%;  Score 35;  DB 23;  Length 858;
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 EVLQEEF 7
                |||||||
Db        255 EVLQEEF 261


RESULT 2
AAB47190
ID   AAB47190 standard; Protein; 860 AA.
XX
AC   AAB47190;
XX
DT   29-JUN-2001  (first entry)
XX
DE   Human DPP8 241Gln-341Ile-515Val 776Ile.
XX
KW   Human; dipeptidyl aminopeptidase; DPP8; prolyl dipeptidase;
KW   dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;
KW   growth hormone deficiency; glucose level; intestinal regeneration;
KW   non-insulin dependent diabetes mellitus; glucose intolerance;
KW   immunosuppression.
XX
OS   Homo sapiens.
XX
PN   WO200119866-A1.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-AU01085.
XX
PR   10-SEP-1999;   99AU-0002762.
PR   28-FEB-2000;   2000AU-0005789.
XX
PA   (UNSY ) UNIV SYDNEY.
XX
PI   Abbott CA,  Gorrell MD;
XX
DR   WPI; 2001-281520/29.
DR   N-PSDB; AAC85597.
XX
PT   New human dipeptidyl aminopeptidase (DPP8) useful for cleaving
PT   substrates, identifying inhibitors of DPP8 catalytic activity which
PT   have therapeutic uses, and for detecting activated T cells [illegible]
XX
PS   [illegible]
XX
CC   The sequence [illegible]
CC   dipeptidyl aminopeptidase [illegible]
CC   H-Ala-Pro-pNA [illegible]
CC   prolyl [illegible] dipeptidase [illegible]
CC   of hydroxyproline [illegible]
CC   compounds [illegible]
CC   cleaving a substrate [illegible]
CC   involves measuring the [illegible]
CC   level of DPP8 expression [illegible]
CC   RNA in the cell [illegible]
CC   of inhibiting the [illegible]
CC   identified as inhibiting [illegible]
CC   treating diarrhoea [illegible]
CC   in non insulin dependent [illegible]
CC   involving disease [illegible]
CC   as immunosuppressants [illegible]
XX
SQ   Sequence   [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy   [illegible]
Db   [illegible]


RESULT 3
[illegible]
ID   [illegible]
XX
AC   [illegible]
XX
DT   [illegible]
XX
DE   Human [illegible]
XX
KW   Human; [illegible] IV-related [illegible]
KW   DPPIV; [illegible]
KW   diabetes; infertility; [illegible]
KW   heart [illegible]
KW   [illegible]
KW   dyskinesia; [illegible]
KW   metabolic disorder.
XX
OS   Homo sapiens.
XX
PN   [illegible]
XX
PD   [illegible]
XX
PF   [illegible]
XX
PR   [illegible]
XX
PA   [illegible]
XX
PI   [illegible]
XX
DR   WPI; [illegible]
DR   N-PSDB; [illegible]
XX
PT   New dipeptidyl peptidase [illegible]
PT   the proteins, [illegible]
PT   viral infections, [illegible]
PT
XX
PS   Disclosure; [illegible]
XX
CC   The present invention [illegible]
```

[illegible]

ALIGNMENTS

[illegible]

XX

KW [illegible] peptidase IV [illegible]
KW [illegible] HIV [illegible]
KW [illegible] obesity; [illegible] Parkinson's disease; [illegible]
KW heart failure; [illegible] retention; osteoporosis; [illegible]
KW ulcer; [illegible] psychotic disorder; neurological disorder;
KW dyskinesia; [illegible] disorder; inflammatory disorder;
KW metabolic [illegible]
XX

[illegible]

```
A;Experimental source: [illegible]
C;Genetics:
A;Note: [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]; Pred. No. [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy [illegible] IVLQERF [illegible]

Db [illegible]


RESULT 11
AB[illegible]
phage protein homolog lmo2[illegible] [imported] - Listeria monocytogenes (strain [illegible])
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AB1[illegible]
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, [illegible]
[illegible] Harris, B.; [illegible]
[illegible] Jones, L.M.; Kunst, F.
Science 294, [illegible] 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; [illegible]
[illegible] Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; [illegible] Wehland, J.
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21[illegible]; PMID:[illegible]
A;Accession: AB1[illegible]
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1[illegible] <GLA>
A;Cross-references: GB:NC_[illegible]; PIDN:[illegible]
A;Experimental source: strain EGD[illegible]
C;Genetics:
A;Gene: lmo2[illegible]

  Query Match              85.7%;  Score 30;  DB 2;  Length 134;
  Best Local Similarity   85.7%;  Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy [illegible] IVLEE [illegible]

Db [illegible] IVLEE [illegible]


RESULT 12
AI1[illegible]
phage protein homolog lin2[illegible] [imported] - Listeria innocua (strain [illegible])
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AI1[illegible]
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, [illegible]
[illegible] Bertrand, [illegible]; Durand, [illegible]
[illegible] Jones, L.M.; Kunst, F.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; [illegible]
[illegible] Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; [illegible] Wehland, J.
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:[illegible]; PMID:[illegible]
A;Accession: AI1[illegible]
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1[illegible] <GLA>
A;Cross-references: GB:AL592022; PIDN:CA[illegible]; GSPDB:GN00[illegible]
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]; Pred. No. [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy [illegible] IVLQER [illegible]
```

[illegible]

```
RESULT 13
[illegible]
[illegible]
C;Species: [illegible]
C;Date: [illegible]
C;Accession: [illegible]
[illegible]
A;Status: preliminary
A;Molecule type: DNA
A;Residues: [illegible]
[illegible]
C;Genetics:
A;Gene: [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy [illegible]

Db [illegible]


RESULT [illegible]
[illegible]
NAD[illegible]
C;Species: [illegible]
A;Variety: [illegible]
C;Date: [illegible]
C;Accession: [illegible]
[illegible]
A;Status: preliminary [illegible]
A;Molecule type: DNA
[illegible]
C;Genetics:
[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy [illegible]

Db [illegible]


RESULT [illegible]
[illegible]
[illegible]
```

[illegible]

FAD-dependent [illegible] malaria parasite [illegible]

[illegible]

A;Title: [illegible] of the human malaria parasite Plasmodium falciparum

[illegible]

C;Superfamily: [illegible] resistance protein

[illegible]

[illegible] Biophys. [illegible]
A;Title: A [illegible] endoplasmic [illegible]
A;Reference number: A4[illegible]; MUID:[illegible]; PMID:[illegible]
A;Accession: A4[illegible]
A;Molecule type: protein
A;Residues: 2[illegible] <MAR>
R;[illegible] Nasi, M.; Moriyama, [illegible] Wada, K.; [illegible] M.
J. Biol. Chem. 2[illegible]
A;Title: Protein degradation by the phosph[illegible]
A;Reference number: A4[illegible]; MUID:[illegible]; PMID:[illegible]
A;Accession: A4[illegible]
A;Molecule type: protein
A;Residues: 2[illegible] <DRA>
R;Morris, [illegible] Fink, [illegible] Pfaff, D.W.
Science 24[illegible] 1477-147[illegible]
A;Title: HIP-7[illegible] a protein induced by estrogen in the brain and [illegible]
A;Reference number: A4[illegible]; MUID:[illegible]; PMID:[illegible]
A;Accession: A4[illegible]
A;Molecule type: protein
A;Residues: 26-43 <MOR>
C;Comment: ER[illegible] may be part of a complex capable of [illegible]
C;Comment: ER[illegible] protease does not appear to be [illegible]
[illegible] inositol 1 phosphate or phosphatidylinositol [illegible] diphosphate.
C;Superfamily: protein disulfide isomerase [illegible]
C;Keywords: duplication; endoplasmic reticulum; [illegible]
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-504/Product: ER60 protease #status experimental <MAT>
F;[illegible]/Domain: thioredoxin homology <TX1>
F;[illegible]/Domain: thioredoxin homology <TX2>
F;[illegible]
F;[illegible]

Query Match [illegible]; Score [illegible]; DB [illegible]; Length 5[illegible];
Best Local Similarity 85.7%; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible] FVL[illegible]

Db [illegible] FVMLEE [illegible]

RESULT [illegible]
S41[illegible]
protein disulfide-isomerase (EC 5.3.4.1) ERp[illegible]
N;Alternate names: ER[illegible] protease; [illegible]
C;Species: Mus musculus (house mouse)
C;Date: [illegible] 1994 #sequence_revision [illegible] #text_change [illegible]
C;Accession: S41[illegible]
R;Mazzarella, R.A.; Marcus, N.; Haugejorden, S.M.; Balcarek, J.M.; [illegible]
Arch. Biochem. Biophys. 308, 454-460, 1994
A;Title: ERp61 is GRP58, a stress-inducible [illegible] endoplasmic reticulum [illegible]
A;Reference number: S41[illegible]; MUID:[illegible]; PMID:[illegible]
A;Accession: S41[illegible]
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-504 <MAZ>
C;Superfamily: protein disulfide isomerase; thioredoxin homology
C;Keywords: endoplasmic reticulum; intramolecular oxidoreductase; [illegible]
F;[illegible]/Domain: thioredoxin homology <TX1>
F;[illegible]/Domain: thioredoxin homology <TX2>

Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
Best Local Similarity 85.7%; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible] FVL[illegible]

Db [illegible] FVML[illegible]

RESULT [illegible]
[illegible]
protein disulfide-isomerase (EC 5.3.4.1) ER[illegible]

[illegible]

RESULT [illegible]
[illegible]
protein disulfide-isomerase [illegible]

[illegible]

[illegible]

SUMMARIES

[illegible]

ALIGNMENTS

[illegible]
protein disulfide [illegible] (EC [illegible]) [illegible] precursor [illegible]
N;Alternate names: [illegible]
[illegible]
[illegible] #text_change [illegible]
[illegible]
[illegible] Hirano, N.; [illegible] Kato, H.; Sakai, K.; Tanaka, [illegible]
[illegible]
[illegible]
A;Molecule type: [illegible]
A;Residues: [illegible]
[illegible]
[illegible]
[illegible]
C;Superfamily: protein disulfide isomerase; thioredoxin homology
C;Keywords: [illegible] endoplasmic reticulum; intramolecular [illegible]
F;[illegible] #status predicted [illegible]
F;[illegible] #status predicted <MAT>
F;[illegible] homology <IX>
F;[illegible] homology <IX>
F;485-488/Region: endoplasmic reticulum retention signal
F;[illegible] #status predicted

Query Match [illegible] Length [illegible]
Best Local Similarity [illegible] Pred. No. [illegible]
Matches [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

[illegible]

[illegible]
C;Species: [illegible] (Norway rat)
C;Date: [illegible] #sequence_revision [illegible] #text_change [illegible]
C;Accession: [illegible]
[illegible] Nature [illegible]
A;Title: [illegible] sequence [illegible]
A;Reference number: [illegible] MUID:[illegible]; PMID:[illegible]
A;Accession: [illegible]
A;Molecule type: [illegible]
A;Residues: [illegible]
A;Cross-references: [illegible] NID:[illegible]
R;Martin, [illegible] A.; Martin, [illegible]

```
CC   -!- FUNCTION: PROTEOLYTICALLY REMOVES THE C-TERMINAL THREE RESIDUES OF
CC       FARNESYLATED A-FACTOR MATING PHEROMONE. [illegible] CLEAVE [illegible]
CC       TERMINAL EXTENSION OF THE PHEROMONE [illegible] IN RAS.
CC   -!- [illegible] BINDS 1 ZINC ION (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein [illegible]
CC       [illegible]
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M48.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]; AA[illegible]47.1; -.
DR   EMBL; [illegible]; AA[illegible]27.1; -.
DR   MEROPS; M48.[illegible]; -.
DR   SGD; S00[illegible]; STE24.
DR   InterPro; IPR[illegible]; Peptidase M48.
DR   InterPro; IPR[illegible]; Zn_M[illegible]peptidase.
DR   Pfam; PF01435; Peptidase_M48; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Hydrolase; Metalloprotease; Zinc; Transmembrane;
KW   Endoplasmic reticulum; Pheromone response.
FT   TRANSMEM     [illegible]   [illegible]   POTENTIAL.
FT   TRANSMEM     [illegible]   [illegible]   POTENTIAL.
FT   TRANSMEM     [illegible]   142           POTENTIAL.
FT   TRANSMEM     [illegible]   [illegible]   POTENTIAL.
FT   TRANSMEM     [illegible]   218           POTENTIAL.
FT   TRANSMEM     [illegible]   327           POTENTIAL.
FT   TRANSMEM     [illegible]   378           POTENTIAL.
FT   METAL        2[illegible]  2[illegible]  ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE     2[illegible]  2[illegible]  BY SIMILARITY.
FT   METAL        [illegible]   [illegible]   ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL        [illegible]   [illegible]   ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE     [illegible]4  394           PROTON DONOR (BY SIMILARITY).
SQ   SEQUENCE   4[illegible] AA;  52421 MW;  [illegible] CRC64;

  Query Match          [illegible]%;  Score 28;  DB [illegible];  Length 453;
  Best Local Similarity   71.4%;  Pred. No. [illegible];
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy        [illegible] IVLLER [illegible]
Db      [illegible] IVLLER [illegible]


RESULT 15
DAL81_YEAST
ID   DAL81_YEAST    STANDARD;      PRT;   970 AA.
AC   P21657;
DT   01-MAY-1991 (Rel. 18, Created)
DT   01-FEB-199[illegible] (Rel. [illegible], Last sequence update)
DT   01-FEB-199[illegible] (Rel. [illegible], Last annotation update)
DE   Transcriptional activator protein DAL81 (Regulatory protein UGA35).
GN   DAL81 OR UGA35 OR DURL OR YIR023W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Coornaert D., Vissers S., Andre B.;
RT   "The pleiotropic UGA35(DURL) regulatory gene of Saccharomyces
RT   cerevisiae: cloning, sequence and identity with the DAL81 gene.";
RL   Gene 97:163-171(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Bricmont P.A., Daugherty J.R., Cooper T.G.;
```

```
[illegible]
RL   [illegible]
RN   [illegible]
RP   [illegible]
RA   [illegible]
CC   -!- [illegible]
CC       [illegible]
CC       [illegible]
CC   -!- [illegible]
CC   -!- [illegible]
CC       [illegible] DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. [illegible]
CC   [illegible]
DR   EMBL; [illegible]
DR   EMBL; [illegible]
DR   EMBL; [illegible]
DR   PIR; [illegible]
DR   PIR; [illegible]
DR   TRANSFAC; [illegible]
DR   SGD; [illegible]
DR   InterPro; [illegible]
DR   Pfam; [illegible]
DR   SMART; [illegible]
DR   PROSITE; [illegible]
DR   PROSITE; [illegible]
KW   [illegible]
KW   [illegible]
FT   DOMAIN       [illegible]
FT   DOMAIN       [illegible]
FT   DOMAIN       [illegible]
FT   DNA_BIND     [illegible]
FT   DOMAIN       [illegible]
FT   CONFLICT     [illegible]
FT   CONFLICT     [illegible]
FT   CONFLICT     [illegible]
FT   CONFLICT     [illegible]
FT   CONFLICT     [illegible]
FT   CONFLICT     [illegible]
SQ   SEQUENCE   [illegible]

  Query Match   [illegible]
  Best Local Similarity   [illegible]
  Matches   [illegible]

Qy   [illegible]

Db   [illegible]


Search completed: [illegible]
Job time : [illegible]
```

```
[illegible]
DE   [illegible] chlorophyll A-C binding protein, chloroplast precursor
DE   [illegible]
[illegible]
RP   SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC   STRAIN [illegible];
RX   MEDLINE [illegible];
RA   [illegible]
RL   Plant Mol. Biol. [illegible](1994).
CC   -!- FUNCTION: THE LIGHT-HARVESTING COMPLEX (LHC) FUNCTIONS AS A LIGHT
CC       RECEPTOR, IT CAPTURES AND DELIVERS EXCITATION ENERGY TO PHOTOSYSTEMS
CC       WITH WHICH IT IS CLOSELY ASSOCIATED. ENERGY IS TRANSFERRED FROM
CC       [illegible]
CC       [illegible]
CC       [illegible]
CC   -!- [illegible] THE LHC COMPLEX [illegible]
CC       [illegible] XANTHIN, CHLOROPHYLL A AND C BOUND NON COVALENTLY BY
CC       [illegible] XANTHIN CHLOROPHYLL [illegible] PROTEINS (FCPS) [illegible]
CC       [illegible]
CC   -!- SUBCELLULAR LOCATION: [illegible] THYLAKOID MEMBRANE [illegible]
CC       [illegible] TRANSMEMBRANE [illegible] MEMBRANE
CC       THAT [illegible] VIA A SIGNAL [illegible]
CC       TRANSLOCATION ACROSS THE THYLAKOID MEMBRANE VIA A TRANSIT PEPTIDE.
CC   -!- INDUCTION: EXPRESSION IS INCREASED [illegible] UNDER CONDITIONS OF
CC       LOW LIGHT.
CC   -!- SIMILARITY: BELONGS TO THE [illegible] FAMILY [illegible] LIGHT HARVESTING
CC       PROTEIN.
CC
CC   [illegible]
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC
DR   EMBL; [illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]
KW   Light-harvesting polypeptide; [illegible]
KW   [illegible]
KW   [illegible] transit peptide.
FT   [illegible]
FT   [illegible] XANTHIN CHLOROPHYLL A [illegible]
FT                                       PROTEIN.
FT   TRANSMEM [illegible]                 POTENTIAL.
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible]

[illegible]

[illegible]

RESULT 1
STE4_YEAST
ID   STE4_YEAST [illegible] AA.
AC   [illegible]
DT   [illegible] (Rel. [illegible] Created)
DT   [illegible] (Rel. [illegible] Last sequence update)
DT   16-[illegible] (Rel. [illegible] Last annotation update)
DE   CAAX prenyl protease [illegible] (EC 3.4.24.[illegible]) (Prenyl protein specific
DE   endoprotease [illegible]) (A factor converting enzyme).
GN   [illegible]
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   [illegible] Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RA   [illegible]
RL   Submitted [illegible] to the EMBL/GenBank/DDBJ databases.
RN   [illegible]
RP   SEQUENCE FROM N.A. [illegible] FUNCTION.
RX   MEDLINE [illegible];
RA   Fujimura-Kamada K., Nouvet F.J., Michaelis S.;
RL   "A novel membrane-associated metalloprotease, Ste24p, is required for
RL   the first step of NH2-terminal processing of the yeast a-factor
RL   precursor.";
RL   [illegible]
RN   [illegible]
RP   FUNCTION.
RX   MEDLINE [illegible];
RA   Boyartchuk V.L., [illegible] M.N., Rine J.;
RT   "Modulation of Ras and a-factor function by carboxyl-terminal
RT   proteolysis.";
RL   Science [illegible].
RN   [illegible]
RP   FUNCTION.
RX   MEDLINE [illegible];
RA   Tam A., Nouvet F.J., Fujimura-Kamada K., Slunt H., Sisodia S.S.,
RA   Michaelis S.;
RT   "Dual roles for Ste24p in yeast a-factor maturation: NH2-terminal
RT   proteolysis and COOH-terminal CAAX processing.";
RL   [illegible]
RN   [illegible]
RP   FUNCTION.
RX   MEDLINE [illegible];
RA   Boyartchuk V.L., Rine J.;
RT   "Roles of prenyl protein proteases in maturation of Saccharomyces
RT   cerevisiae a-factor.";
RL   Genetics [illegible].
RN   [illegible]
RP   FUNCTION.
RX   MEDLINE [illegible];
RA   Schmidt W.K., [illegible] Michaelis S.;
RT   "[illegible] Ste24p-dependent N-terminal proteolytic [illegible]
RT   yeast a-factor [illegible]";
RL   [illegible]
RN   [illegible]
RP   CHARACTERIZATION.
RX   MEDLINE [illegible];
RA   [illegible]
RA   [illegible]
RT   "The CaaX proteases [illegible] Ste24p and Rce1p, have overlapping but distinct
RT   substrate specificities.";
RL   Mol. [illegible]
RN   [illegible]
RP   SUBCELLULAR LOCATION.
RX   MEDLINE [illegible];
RA   Schmidt W.K., [illegible] Fujimura-Kamada K., Michaelis S.;
RT   "Endoplasmic reticulum membrane localization of Rce1p and Ste24p,
RT   yeast proteases involved in carboxyl-terminal CAAX protein processing
RT   and amino-terminal a-factor cleavage.";
RL   Proc. Natl. Acad. Sci. U.S.A. [illegible] (1998).
```

```
CC   [illegible] IN THE MULTI PROTEIN COMPLEX [illegible]
CC       [illegible] AND THE [illegible]
CC       [illegible] A CYLINDER SHAPED
CC       COMPLEX OF FOUR STACKED RINGS, EACH CONTAINING SEVEN SUBUNITS.
CC   [illegible] (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. [illegible]
CC   between the Swiss Institute of Bioinformatics and the EMBL [illegible]
CC   the European Bioinformatics Institute. [illegible]
CC   use by non-profit institutions as long as its content is [illegible]
CC   modified and this statement is not removed. [illegible]
CC   entities requires a license agreement (See http://www.isb-sib.ch/[illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]; BAA[illegible]
DR   EMBL; AF[illegible]; AAB[illegible]
DR   EMBL; [illegible]; AAB[illegible]
DR   EMBL; AK[illegible]; BAB[illegible]
DR   EMBL; [illegible]; AAH[illegible]
DR   Genew; HGNC:[illegible]; PSMC[illegible]
DR   MIM; [illegible]
DR   [illegible]
DR   InterPro; IPR[illegible]; AAA ATPase central.
DR   InterPro; IPR[illegible]; AAA sub.
DR   Pfam; PF[illegible]; AAA; 1.
DR   SMART; SM[illegible]; 26Sp45; 1.
DR   PROSITE; PS[illegible]; AAA; 1.
KW   Proteasome; ATP-binding; Nuclear protein.
FT   NP_BIND    [illegible]    ATP (POTENTIAL).
FT   [illegible]
SQ   SEQUENCE   [illegible] AA;  441[illegible] MW;  [illegible] CRC64;

  Query Match             [illegible];  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   71.4%;  Pred. No. [illegible];
  Matches    5;  Conservative    [illegible];  Mismatches    [illegible];  Indels    [illegible];  Gaps

Qy          1 EVLQEEF 7
              [illegible]
Db        361 EVVQEDF 367


RESULT 11
VE7_HPV49
ID   VE7_HPV49      STANDARD;      PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] 1994 (Rel. [illegible], Created)
DT   [illegible] 1994 (Rel. [illegible], Last sequence update)
DT   [illegible] (Rel. [illegible], Last annotation update)
DE   E7 protein.
GN   E7.
OS   Human papillomavirus type 49.
OC   Viruses; dsDNA viruses, no RNA stage; [illegible]
OC   Papillomavirus.
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible]
RA   [illegible]
RT   "Primer-directed sequencing of human papillomavirus types.";
RL   [illegible]
CC   -!- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANSACTIVATING
CC       ACTIVITIES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. [illegible]
CC   between the Swiss Institute of Bioinformatics and the EMBL [illegible]
CC   the European Bioinformatics Institute. [illegible]
CC   use by non-profit institutions as long as its content is [illegible]
CC   modified and this statement is not removed. [illegible]
CC   entities requires a license agreement (See http://www.isb-sib.ch/[illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X7[illegible]; CAA[illegible]
```

```
DR   [illegible]
DR   [illegible]
DR   [illegible]
KW   [illegible]
KW   [illegible]
FT   [illegible]
FT   [illegible]
SQ   SEQUENCE   [illegible]

  Query Match   [illegible]
  Best Local Similarity   [illegible]
  Matches   [illegible]

Qy   [illegible]

Db   [illegible]


RESULT [illegible]
[illegible]
ID   [illegible]
AC   [illegible]
DT   [illegible]
DT   [illegible]
DT   [illegible]
DE   [illegible]
GN   [illegible]
OS   [illegible]
OC   [illegible]
OC   [illegible]
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=[illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible]
RA   [illegible]
RA   Kerlavage A.R., [illegible]
RA   McKenney K., [illegible]
RA   Scott J.L., [illegible] Kelley J.M.
RA   Weidman J.F., [illegible]
RA   [illegible]
RA   [illegible]
RA   [illegible]
RA   Venter J.C.;
RT   "Whole-genome [illegible]
RT   [illegible]"
RL   [illegible]
CC   [illegible]
CC   [illegible]
CC   --------------------------------------------------------------------------
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]
KW   [illegible]
SQ   SEQUENCE   [illegible]

  Query Match   [illegible]
  Best Local Similarity   [illegible]
  Matches   [illegible]

Qy   [illegible]

Db   [illegible]
```

```
DX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   [illegible] Brain;
RX   MEDLINE [illegible]; PubMed [illegible];
RA   Monyer H., Sprengel R., Schoepfer R., Herb A., Higuchi M., Lomeli H.,
RA   Burnashev N., Sakmann B., Seeburg P.H.;
RT   "Heteromeric NMDA receptors: molecular and functional distinction of
RT   subtypes.";
RL   Science 256:1217-1221(1992).
RN   [2]
RP   REVISIONS [illegible] AND [illegible].
RA   Monyer H., Sprengel R., Schoepfer R., Herb A., Higuchi M., Lomeli H.,
RA   Burnashev N., Sakmann B., Seeburg P.H.;
RL   Submitted ([illegible]) to [illegible].
CC   -!- FUNCTION: NMDA RECEPTOR SUBTYPE OF GLUTAMATE-GATED ION CHANNELS
CC       POSSESSES HIGH CALCIUM PERMEABILITY AND VOLTAGE DEPENDENT
CC       SENSITIVITY TO MAGNESIUM AND IS MEDIATED BY GLYCINE.
CC   -!- SUBUNIT: HETERODIMER OF AN EPSILON SUBUNIT AND A ZETA SUBUNIT.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://[illegible]/
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; M91561; AAC[illegible].1; -.
DR   HSSP; P[illegible]; 1GR2.
DR   InterPro; IPR001320; Ion_glu_receptor.
DR   InterPro; IPR001311; SBP/glu_receptor.
DR   Pfam; PF00060; Lig_chan; 1.
DR   PRODOM; PD000500; Ion_glu_receptor; 1.
DR   SMART; SM00079; PBPe; 1.
KW   Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW   Transmembrane; Calcium; Magnesium.
FT   SIGNAL        1     [illegible]       POTENTIAL.
FT   CHAIN        [illegible]  1464       GLUTAMATE [NMDA] RECEPTOR SUBUNIT
FT                                EPSILON 1.
FT   DOMAIN       [illegible]  [illegible]  EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     [illegible]  [illegible]  1 (POTENTIAL)
FT   TRANSMEM     [illegible]  [illegible]  2 (POTENTIAL).
FT   TRANSMEM     [illegible]  [illegible]  3 (POTENTIAL).
FT   TRANSMEM     [illegible]  [illegible]  4 (POTENTIAL).
FT   SITE         [illegible]  [illegible]  FUNCTIONAL DETERMINANT OF NMDA
FT                                RECEPTORS (BY SIMILARITY).
FT   CARBOHYD     [illegible]  [illegible]  N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     [illegible]  [illegible]  N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     44[illegible] 443       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     444   444       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     [illegible]  [illegible]  N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   1464 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score [illegible];  DB [illegible];  Length 1464;
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy        [illegible] EVLQEE [illegible]
               [illegible]
Db    [illegible] EMLQEE [illegible]


RESULT [illegible]
PRSX_HUMAN
ID   PRSX_HUMAN     STANDARD;      PRT;   389 AA.
AC   [illegible];
DT   [illegible] (Rel. [illegible], Created)
DT   [illegible] NOV [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] JUN 2002 (Rel. [illegible], Last annotation update)
DE   26S protease regulatory subunit S10B (Proteasome subunit p42) (p44)
DE   (Conserved ATPase domain protein 44) (CADp44)
GN   PSMC6 OR [illegible].
OS   Homo sapiens (Human),
OS   Mus musculus (Mouse), and
OS   Spermophilus tridecemlineatus (Thirteen-lined ground squirrel).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=[illegible], [illegible], 43179;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Human; TISSUE=Aorta;
RX   MEDLINE=96211[illegible]; PubMed=8674546;
RA   Fujiwara T., Watanabe T.K., Tanaka K., Slaughter C.A.,
RA   DeMartino G.N.;
RT   "[illegible] proteasome regulatory
RT   proteins, reveals a novel member of the AAA protein family.";
RL   FEBS Lett. [illegible](1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Human;
RA   Li Y., [illegible] K.;
RT   "Human SUG[illegible].";
RL   Submitted (MAY [illegible]) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Human; TISSUE=Urinary bladder;
RA   Strausberg R.;
RL   Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   PARTIAL SEQUENCE.
RC   SPECIES=Human;
RX   MEDLINE=[illegible]; PubMed=862[illegible];
RA   DeMartino G.N., [illegible] R.J., Moomaw C.R., Strong A.A., Song X.,
RA   Hisamatsu H., Tanaka K., Slaughter C.A.;
RT   "Identification, purification, and characterization of a PA700-
RT   dependent activator of the proteasome.";
RL   J. Biol. Chem. [illegible]:3112-3118(1996).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Mouse; STRAIN=C57BL/6J;
RX   MEDLINE=21085[illegible]; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., [illegible] M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature [illegible]([illegible]).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.tridecemlineatus;
RX   MEDLINE=97126[illegible]; PubMed=8973[illegible];
RA   Bauer V.W., Swaffield J.C., Johnston S.A., Andrews M.T.;
RT   "CADp44: a novel regulatory subunit of the 26S proteasome and the
RT   mammalian homolog of yeast Sug2p.";
RL   Gene [illegible]([illegible]).
CC   -!- FUNCTION: THE 26S PROTEASE IS INVOLVED IN THE ATP-DEPENDENT
CC       DEGRADATION OF UBIQUITINATED PROTEINS. THE REGULATORY (OR ATPASE)
CC       COMPLEX CONFERS ATP DEPENDENCY AND SUBSTRATE SPECIFICITY TO THE
CC       26S COMPLEX (BY SIMILARITY).
```

```
RT   [illegible] aspartate recept[illegible]";
RL   [illegible] Pharmacol. Exp. Ther. [illegible]
CC   -!- FUNCTION: NMDA RECEPTOR SUBTYPE OF GLUTAMATE-GATED ION CHANNELS
CC       POSSESSES HIGH CALCIUM PERMEABILITY AND VOLTAGE-DEPENDENT
CC       SENSITIVITY TO MAGNESIUM AND IS MODULATED BY GLYCINE.
CC   -!- SUBUNIT: HETEROMIMER OF AN EPSILON SUBUNIT AND A ZETA SUBUNIT.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE LIGAND-GATED ION CHANNEL FAMILY.
CC   --------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. [illegible]
CC   entities requires a license agreement (See http://www.isb-sib.ch/[illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------
DR   EMBL; [illegible]; AA[illegible].1; -.
DR   EMBL; [illegible]; AAB4[illegible]2.1; -.
DR   HSSP; [illegible]; [illegible].
DR   Genew; HGNC:4[illegible]; GRIN2A.
DR   MIM; [illegible]; -.
DR   InterPro; [illegible]; [illegible] receptor.
DR   InterPro; IPR[illegible]; SBP/[illegible] receptor.
DR   Pfam; [illegible]; [illegible]_chan; 1.
DR   Pfam; [illegible]; Lig_chan_recept [illegible]; 1.
DR   SMART; SM[illegible]; [illegible]; 1.
KW   Receptor; Postsynaptic membrane; [illegible]
KW   Transmembrane; Calcium; Magnesium; Polymorphism.
FT   SIGNAL        [illegible]     22       POTENTIAL.
FT   CHAIN         [illegible]   1464       GLUTAMATE [NMDA] RECEPTOR SUBUNIT
FT                                          EPSILON 1.
FT   DOMAIN        [illegible]  [illegible]  EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      [illegible]  [illegible]  1 (POTENTIAL).
FT   TRANSMEM      [illegible]  [illegible]  2 (POTENTIAL).
FT   TRANSMEM      [illegible]  [illegible]  3 (POTENTIAL).
FT   TRANSMEM      [illegible]  [illegible]  4 (POTENTIAL).
FT   SITE          [illegible]    614       FUNCTIONAL DETERMINANT OF NMDA
FT                                          RECEPTORS (BY SIMILARITY).
FT   CARBOHYD      [illegible]     75       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD      [illegible]  [illegible]  N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD      [illegible]  [illegible]  N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD      [illegible]  [illegible]  N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD      [illegible]    444       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD      [illegible]  [illegible]  N-LINKED (GLCNAC...) (POTENTIAL).
FT   VARIANT       [illegible]  [illegible]  K -> [illegible].
FT                                          /FTId=VAR_[illegible].
SQ   SEQUENCE   1464 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             85.7%;  Score [illegible];  DB [illegible];  Length 1464;
  Best Local Similarity   71.4%;  Pred. No. [illegible];
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy   [illegible] [illegible]VLEH[illegible]
Db   [illegible] [illegible]MLEH[illegible]


RESULT 8
NMEI_MOUSE
ID   NMEI_MOUSE     STANDARD;      PRT;  1464 AA.
AC   [illegible];
DT   [illegible]-JUN-1994 (Rel. 29, Created)
DT   [illegible]-JUN-1994 (Rel. 29, Last sequence update)
DT   [illegible]-JUN-[illegible] (Rel. [illegible], Last annotation update)
DE   Glutamate [NMDA] receptor subunit epsilon 1 precursor (N-methyl
DE   D-aspartate receptor subtype 2A) (NR2A) (NMDAR2A).
GN   GRIN2A.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=[illegible];
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible];
RA   [illegible]
RT   [illegible]
RL   [illegible]
CC   -!- FUNCTION: NMDA RECEPTOR SUBTYPE OF GLUTAMATE-GATED ION CHANNELS
CC       POSSESSES HIGH CALCIUM PERMEABILITY AND VOLTAGE-DEPENDENT
CC       SENSITIVITY TO MAGNESIUM AND IS MODULATED BY GLYCINE.
CC   -!- SUBUNIT: HETEROMIMER OF AN EPSILON SUBUNIT AND A ZETA SUBUNIT.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE LIGAND-GATED ION CHANNEL FAMILY.
CC   [illegible]
DR   [illegible]
KW   [illegible]
FT   SIGNAL        [illegible]
FT   CHAIN         [illegible]
FT   DOMAIN        [illegible]
FT   TRANSMEM      [illegible]
FT   TRANSMEM      [illegible]
FT   TRANSMEM      [illegible]
FT   TRANSMEM      [illegible]
FT   SITE          [illegible]
FT   CARBOHYD      [illegible]
SQ   SEQUENCE   1464 AA;  [illegible]

  Query Match  [illegible]
  Best Local Similarity  [illegible]
  Matches  [illegible]

Qy   [illegible]
Db   [illegible]

RESULT 9
NMEI_[illegible]
ID   NME[illegible]
AC   [illegible]
DT   [illegible]
DE   Glutamate [NMDA] receptor subunit [illegible]
DE   [illegible] (NMDAR2A).
GN   GRIN2A.
OS   [illegible]
OC   [illegible]
OC   [illegible]
```

```
[illegible]
KW   [illegible] pathway [illegible]
FT   ACT_SITE [illegible] BY SIMILARITY.
FT   ACT_SITE [illegible] BY SIMILARITY.
FT   ACT_SITE [illegible] BY SIMILARITY.
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Mismatches [illegible] Indels [illegible]

[illegible]

[illegible]
[illegible] HUMAN
ID   [illegible]_HUMAN      STANDARD;      PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] MAY [illegible] (Rel. [illegible], Created)
DT   [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] (Rel. [illegible], Last annotation update)
DE   [illegible] (EC [illegible]) ([illegible]
DE   [illegible]) ([illegible] specific protease [illegible])
DE   ([illegible]) ([illegible]).
GN   [illegible]
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. ([illegible] AND [illegible]).
RC   TISSUE=Fetal brain;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible], Martinez [illegible]
RA   [illegible]
RT   "[illegible]
RT   [illegible]";
RL   [illegible]
RN   [2]
RP   SEQUENCE FROM N.A. ([illegible]).
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible], Lasek [illegible], Baird L.A., Tamiton M., [illegible]
RA   [illegible] K., Nisbet [illegible]
RT   "[illegible]
RT   [illegible]
RT   [illegible]."
RL   [illegible]
RN   [3]
RP   SEQUENCE [illegible] FROM N.A. ([illegible]
RA   [illegible]
RA   Nakajima [illegible]
RA   Matsumura [illegible], Kawai [illegible], Yamamoto [illegible], Wakamatsu A.,
RA   Nakamura Y., [illegible], Sasaki N.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (APR-[illegible]) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE [illegible] PATHWAY
CC       [illegible] BY SIMILARITY.
CC   -!- CATALYTIC ACTIVITY: [illegible]
CC       [illegible]
CC   -!- ALTERNATIVE PRODUCTS: [illegible] FORMS; A [illegible] FORM (SHOWN HERE) AND A
CC       SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: FOUND IN [illegible] AND FETAL [illegible]
CC       EXPRESSION [illegible] MODERATELY [illegible]
CC       [illegible]
CC   [illegible]
CC       [illegible]
CC   -!- SIMILARITY: [illegible] FAMILY. [illegible]
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]
DR   EMBL; [illegible]
DR   EMBL; [illegible]
DR   MEROPS; [illegible]
DR   Genew; [illegible]
DR   MIM; [illegible]
DR   InterPro; [illegible]
DR   InterPro; [illegible]
DR   InterPro; [illegible]
DR   Pfam; [illegible]
DR   Pfam; [illegible]
DR   Pfam; [illegible]
DR   PROSITE; [illegible]
DR   PROSITE; [illegible]
DR   PROSITE; [illegible]
KW   [illegible]
KW   Alternative splicing.
FT   ACT_SITE [illegible] BY SIMILARITY.
FT   ACT_SITE [illegible] BY SIMILARITY.
FT   ACT_SITE [illegible] BY SIMILARITY.
FT   VARSPLIC [illegible] MISSING (IN SHORT ISOFORM).
FT   CONFLICT [illegible] (IN REF. [illegible]).
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Mismatches [illegible] Indels [illegible]

[illegible]

[illegible]
NMEL HUMAN
ID   NMEL_HUMAN      STANDARD;      PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] (Rel. [illegible], Created)
DT   16-OCT-2001 (Rel. [illegible], Last sequence update)
DT   [illegible] JUN 200[illegible] (Rel. [illegible], Last annotation update)
DE   [illegible] (NMEA) [illegible] endopeptidase [illegible] precursor (N[illegible]
DE   [illegible] (NEPA) (NMEARA) (hNEPA).
GN   [illegible]
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. AND VARIANT [illegible].
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible], Fantasko [illegible], Kumf [illegible];
RT   "Human N[illegible] mediator [illegible]
RT   and sequence [illegible] cDNA and primary structure of the protein."
RL   [illegible]
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=[illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible]
RA   [illegible] Volkeveeb [illegible]
RT   "[illegible] characterization [illegible] human [illegible]
```

```
RN   [illegible]
RP   [illegible]
RA   [illegible]
RT   "Role of novel microsomal cysteine proteases.";
RL   [illegible]
RN   [illegible]
RP   PARTIAL [illegible]
RC   TISSUE=Liver;
RX   MEDLINE [illegible]; PubMed [illegible];
RA   Martin [illegible]
RA   [illegible]
RT   "A [illegible] endoplasmic
RT   reticulum protein that is highly homologous to
RT   phosphatidylinositol-specific phospholipase C alpha [illegible]
RT   activity.";
RL   [illegible]
RN   [4]
RP   SEQUENCE OF [illegible]
RC   TISSUE=Brain, and Pituitary;
RX   MEDLINE [illegible]; PubMed [illegible];
RA   [illegible]
RT   "[illegible] a protein induced by estrogen in the brain and [illegible] in the
RT   pituitary.";
RL   [illegible]
RN   [illegible]
RP   SEQUENCE OF [illegible] AND [illegible]
RC   TISSUE=Liver;
RX   MEDLINE [illegible]; PubMed [illegible];
RA   [illegible]
RT   "Purification and characterization of [illegible]
RT   disulfide oxidoreductase from rat hepatic microsomes. Relationship [illegible]
RT   this isozyme [illegible] cytosolic phosphatidylinositol-specific phospholipase
RT   [illegible]";
RL   J. Biol. Chem. [illegible]
RN   [illegible]
RP   SEQUENCE OF [illegible] AND [illegible]
RX   MEDLINE [illegible]; PubMed [illegible];
RA   [illegible] M., [illegible]
RT   "Protein degradation by the phosphoinositide-specific phospholipase
RT   C-alpha family from rat liver endoplasmic reticulum.";
RL   J. Biol. Chem. [illegible](1992).
RN   [illegible]
RP   INHIBITION BY [illegible]
RX   MEDLINE [illegible]; PubMed [illegible];
RA   [illegible]
RT   "Inhibition by acidic phospholipids of protein degradation by [illegible]
RT   protease [illegible] cysteine protease of endoplasmic reticulum.";
RL   FEBS [illegible](1992).
CC   -!- CATALYTIC ACTIVITY: Rearrangement of both intrachain and
CC       interchain disulfide bonds in proteins [illegible]
CC       structures.
CC   -!- ENZYME REGULATION: SEEMS TO BE INHIBITED BY [illegible]
CC   -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC   -!- SIMILARITY: CONTAINS 2 [illegible] REDOXIN DOMAINS.
CC   -!- [illegible] WAS ORIGINALLY [illegible]
CC       [illegible]
CC       [illegible] ALPHA; THEN WAS THOUGHT [illegible] AND [illegible]
CC       PROTEASE.
CC   [illegible]
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   [illegible]
DR   EMBL; X12[illegible]; CAA[illegible]
DR   EMBL; [illegible]; CAA[illegible]
DR   PIR; A[illegible]; A[illegible]
DR   PIR; A[illegible]; A[illegible]
DR   HSSP; [illegible]; [illegible]
DR   MEROPS; [illegible]
```

```
[illegible]
```

KW Redox-active center; Isomerase; Endoplasmic reticulum; Repeat; Signal;

FT SIGNAL

FT CHAIN PROTEIN DISULFIDE ISOMERASE A

FT DISULFID REDOX ACTIVE (BY SIMILARITY).

FT DISULFID REDOX ACTIVE (BY SIMILARITY).

FT SITE PREVENT SECRETION FROM ER

FT (BY SIMILARITY).

SQ SEQUENCE AA; MW;

```
DR   [illegible]
DR   [illegible]
DR   PRINTS; PR0[illegible]; [illegible] XIN.
DR   PFAM; PF0[illegible]; p4H_dom; 2.
DR   PFAM; PF0[illegible]; [illegible] 1.
DR   PROSITE; PS[illegible]; [illegible] XIN; 2.
KW   [illegible]
KW   Phosphorylation.
FT   SIGNAL       1     24
FT   CHAIN        [illegible]  [illegible]   PROTEIN [illegible] IS [illegible]ASE A.
FT   DISULFID     [illegible]  [illegible]   REDOX-ACTIVE (BY SIMILARITY).
FT   DISULFID     4[illegible]  4[illegible]   REDOX-ACTIVE (BY SIMILARITY).
FT   SITE         5[illegible]  504   PREVENT SECRETION FROM ER
FT                            (BY SIMILARITY).
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score [illegible];  [illegible]  Length [illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy    [illegible] EVLEE [illegible]
         [illegible]
Db    [illegible] EVMEEE [illegible]


RESULT [illegible]
PDA[illegible]_BOVIN
ID   PDA[illegible]_BOVIN    STANDARD;    PRT;   [illegible] AA.
AC   P[illegible];
DT   [illegible] FEB 1995 (Rel. 31, Created)
DT   [illegible] FEB 1995 (Rel. 31, Last sequence update)
DT   [illegible] 2001 (Rel. 40, Last annotation update)
DE   Protein disulfide isomerase A[illegible] precursor (EC 5.3.4.1) (Disulfide
DE   isomerase ER[illegible]) (ERP[illegible]) (58 kDa microsomal protein) (P58) (ERP57,
GN   [illegible].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Hirano N., Shibasaki F., Sakai [illegible], [illegible], Nishida [illegible],
RA   Yazaki Y., Takenawa T., Hirai H.;
RL   "Molecular cloning of the human glucose-regulated protein
RL   ERp57/GRP58, a thiol-dependent reductase. [illegible]
RL   [illegible] by TCR and inductive expression [illegible]
RL   transformation.";
RL   Eur. J. Biochem. 234:[illegible](199[illegible]).
CC   -!- CATALYTIC ACTIVITY: Rearrangement of [illegible] and
CC       interchain disulfide bonds in proteins [illegible]
CC       structures.
CC   -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (By similarity).
CC   -!- SIMILARITY: CONTAINS 2 THIOREDOXIN DOMAINS.
CC   -!- CAUTION: WAS ORIGINALLY [illegible]
CC       [illegible] PHOSPHOLIPASE [illegible] (PI-PLC ALPHA).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/[illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]; BAA[illegible]; -.
DR   HSSP; [illegible]; [illegible].
DR   [illegible]
DR   [illegible]
DR   PRINTS; PR0[illegible]; [illegible] XIN.
DR   PFAM; PF0[illegible]; p4H_dom; 2.
DR   PFAM; PF0[illegible]; [illegible]
```

```
DR   [illegible]
KW   [illegible]
FT   [illegible]
FT   CHAIN    [illegible]
FT   [illegible]
FT   [illegible]
FT   SITE     [illegible]
FT   [illegible]
SQ   SEQUENCE   [illegible]

  Query Match   [illegible]
  Best Local Similarity   [illegible]
  Matches   [illegible]

Qy   [illegible]

Db   [illegible] EVME[illegible]


RESULT [illegible]
PDA[illegible]_HUMAN
ID   PDA[illegible]_HUMAN    STANDARD;    PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible]
DT   [illegible]
DT   [illegible]
DE   Protein disulfide isomerase A[illegible]
DE   [illegible]
GN   [illegible]
OS   Homo sapiens (Human).
OC   [illegible]
OC   Mammalia; [illegible]
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]
RA   [illegible]
RA   [illegible]
RL   "Molecular [illegible]
RL   [illegible]
RL   [illegible]
RL   [illegible]
RN   [2]
RP   SEQUENCE FROM N.A.
RC   [illegible]
RA   [illegible]
RL   Submitted [illegible]
RN   [3]
RP   SEQUENCE FROM N.A.
RA   [illegible]
RL   Submitted [illegible]
RN   [4]
RP   SEQUENCE FROM N.A.
RC   [illegible]
RX   MEDLINE=[illegible]
RA   [illegible]
RA   [illegible]
RL   "[illegible]
RL   [illegible]
RL   [illegible]
RL   [illegible]
RL   [illegible]
RN   [5]
RP   SEQUENCE [illegible]
RX   MEDLINE=[illegible]
RA   [illegible]
RA   [illegible]
RL   "[illegible]
RL   [illegible]
RL   [illegible]
RL   [illegible]
RN   [illegible]
```

[illegible]

ALIGNMENTS

[illegible]

```
[illegible]
FLA  MOUSE
ID   [illegible] AA.
AC   [illegible]
DT   [illegible] (Rel. [illegible], Created)
DT   [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] (Rel. [illegible], Last annotation update)
DE   Protein disulfide isomerase A[illegible] precursor (EC [illegible]) ([illegible]
DE   isomerase [illegible]) ([illegible]) ([illegible] microsomal protein) ([illegible]) ([illegible]
GN   [illegible]
OS   Mus [illegible]
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; M[illegible]
OX   [illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE [illegible]
RA   [illegible]
RT   "Expression [illegible] lipase [illegible] by murine [illegible]";
RL   J. Immunol. [illegible] ([illegible]).
RN   [2]
RP   SEQUENCE [illegible]
RC   TISSUE=[illegible]
RX   MEDLINE [illegible]
RA   Merrick B.A., [illegible] Wichter [illegible] Selkirk [illegible]
RT   "Separation and sequencing of familiar and novel murine proteins
RT   using preparative two-dimensional gel electrophoresis.";
RL   [illegible]
RN   [3]
RP   SEQUENCE [illegible] AND PHOSPHORYLATION.
RC   [illegible]
RA   Merrick B.A., [illegible] Selkirk [illegible]
RT   "Identification [illegible] two isoforms of phospholipase C alpha [illegible]
RT   dividing murine [illegible] by protein microsequencing.";
RL   [illegible]
CC   -!- [illegible] Rearrangement of both intrachain and
CC       interchain disulfide bonds in proteins to form the native
CC       [illegible]
CC   -!- [illegible]
CC   -!- [illegible]
CC   -!- SIMILARITY: CONTAINS 2 THIOREDOXIN DOMAINS.
CC   -!- [illegible]
CC       [illegible] (PHOSPHOLIPASE C ALPHA).
CC   [illegible]
CC   [illegible] copyright. It is created through a collaboration
CC   between [illegible] and the [illegible]
CC   the [illegible] Institute. There are no restrictions on its
CC   use [illegible] institutions as long as its content is in no way
CC   modified [illegible] not removed. [illegible]
CC   entities [illegible] license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   [illegible]
DR   EMBL; [illegible]
DR   [illegible]
DR   SWISS-2DPAGE; [illegible] MOUSE.
DR   MGD; MGI:[illegible]
```

```
DR   [illegible]
DR   [illegible] pdi_dom; [illegible]
DR   [illegible]
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible]

Qy   [illegible]
Db   [illegible] EVMQEFF [illegible]

RESULT [illegible]
[illegible]
ID   [illegible] PRELIMINARY; [illegible] AA.
AC   [illegible]
DT   [illegible]-MAR-2[illegible] (TrEMBLrel. 2[illegible], Created)
DT   [illegible]-MAR-2[illegible] (TrEMBLrel. 2[illegible], Last sequence update)
DT   [illegible]-JUN-2002 (TrEMBLrel. 2[illegible], Last annotation update)
DE   Dipeptidyl peptidase [illegible].
GN   DPP[illegible].
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=[illegible]
RN   [illegible]
RP   SEQUENCE FROM N.A.
RA   [illegible], Wagtmann N.;
RT   "Identification and characterization of a novel member of the
RT   dipeptidyl peptidase IV-related family.";
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF452102; AAL47179.1; -.
DR   InterPro; IPR002469; DPPIV_N_term.
DR   InterPro; IPR[illegible]; Peptidase_S9.
DR   InterPro; IPR[illegible]; Pro_amidase.
DR   InterPro; IPR[illegible]; Ser_estrs_site.
DR   Pfam; PF[illegible]; DPPIV_N_term; 1.
DR   Pfam; PF[illegible]; Peptidase_S9; 1.
DR   PRINTS; PR[illegible]; PROAMNPTASE.
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible]

  Query Match [illegible]; Score [illegible]; DB [illegible]; Length 863;
  Best Local Similarity [illegible]; Pred. No. [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible]

Qy   [illegible] EVLEEF [illegible]
Db   [illegible] EVLEEF [illegible]

RESULT [illegible]
[illegible]
ID   [illegible] PRELIMINARY; [illegible] AA.
AC   [illegible]
DT   [illegible]-MAY-[illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible]-MAY-[illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible]-MAY-[illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   [illegible]
GN   [illegible]
OS   [illegible]
OC   [illegible]
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible]
RA   [illegible]
RA   [illegible]
RA   [illegible]
RA   [illegible]
RT   [illegible]
```

```
DR   [illegible]
SQ   SEQUENCE [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy   [illegible]
Db   [illegible]

RESULT [illegible]
[illegible]
ID   [illegible]
AC   [illegible]
DT   [illegible]
DT   [illegible]
DT   [illegible]
DE   [illegible]
GN   [illegible]
OS   [illegible]
OC   [illegible]
OC   [illegible]
OX   NCBI_TaxID=[illegible]
RN   [illegible]
RP   SEQUENCE FROM N.A.
RP   STRAIN [illegible]
RX   MEDLINE [illegible]
RA   [illegible]
RA   [illegible]
RA   [illegible]
RA   [illegible]
RT   [illegible]
RT   [illegible]
RL   [illegible]
DR   [illegible]
SQ   SEQUENCE [illegible] AA; [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy   [illegible]
Db   [illegible]

RESULT [illegible]
[illegible]
ID   [illegible]
AC   [illegible]
DT   [illegible]
DT   [illegible]
DT   [illegible]
DE   [illegible]
GN   [illegible]
OS   Homo sapiens [illegible]
OC   Eukaryota; Metazoa; [illegible]
OC   Mammalia; [illegible]
OX   NCBI_TaxID=[illegible]
RN   [illegible]
RP   SEQUENCE FROM N.A.
RA   [illegible]
RL   [illegible]
DR   [illegible]
```

[illegible]

```
ID   [illegible]    PRELIMINARY;    PRT;    [illegible] AA.
AC   [illegible]
DT   [illegible] NOV [illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] NOV [illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] NOV [illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   [illegible] (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; [illegible]
OX   NCBI_TaxID=[illegible];
RN   [illegible]
RP   SEQUENCE FROM N.A.
RA   [illegible]
RL   Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
[illegible]

ID   [illegible]    PRELIMINARY;    PRT;    [illegible] AA.
AC   [illegible]
DT   [illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   [illegible]
OS   [illegible] (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; [illegible]; Sciurognathi; Muridae; Cricetinae;
OC   [illegible]
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RA   [illegible]
RT   "[illegible] the Chinese hamster [illegible]";
RL   Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; [illegible]
DR   InterPro; [illegible]
DR   Pfam; [illegible]
```

```
RESULT 2
[illegible]
ID   [illegible]     PRELIMINARY;     PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   CDNA FLJ[illegible] fis, clone [illegible].
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kawakami [illegible]
RA   Nakajima Y., [illegible]
RA   Nishi [illegible], Tanaka [illegible], Nakamura [illegible]
RT   "NEDO human cDNA sequencing project.";
RL   Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK[illegible]; BAA91[illegible].
DR   MEROPS; S09.[illegible];
SQ   SEQUENCE   [illegible] AA;  51[illegible] MW;  [illegible] CRC64;

  Query Match   [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]

Qy   [illegible] EVLDE [illegible]
Db   [illegible] EVLDED [illegible]


RESULT 3
[illegible]
ID   [illegible]     PRELIMINARY;     PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] DEC 2001 (TrEMBLrel. 19, Created)
DT   [illegible] DEC 2001 (TrEMBLrel. 19, Last sequence update)
DT   [illegible]-JUN 2002 (TrEMBLrel. 21, Last annotation update)
DE   CDNA FLJ14920 fis, clone PLACE1007416, weakly similar to dipeptidyl
DE   peptidase IV (EC 3.4.14.5).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RA   [illegible], Hayashi K., [illegible], Suzuki Y.,
RA   Nishikawa [illegible], Nagai K., Sugano S., Shirai [illegible] A., [illegible] H.,
RA   Wakatsuki M., Hosoiri [illegible], Kaku Y., Kodaira H., [illegible] H., [illegible]
RA   Takahashi M., Shiba Y., Ishii [illegible], Murakawa K., [illegible] Y., Takiguchi
RA   Watanabe S., Kimura K., Murakami K., [illegible], Kawai Y., [illegible]
RA   Yamamoto [illegible], Wakamatsu A., Nakamura Y., [illegible], Masuho Y.,
RA   [illegible] K., [illegible]
RT   "NEDO human cDNA sequencing [illegible].";
RL   Submitted (MAY-[illegible]) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK02[illegible]; BAB[illegible].
DR   MEROPS; S09.[illegible];
SQ   SEQUENCE   [illegible] AA;  72[illegible] MW;  [illegible] CRC64;

  Query Match   [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]

Qy   [illegible] EVLDE [illegible]
Db   [illegible] EVLDED [illegible]


RESULT 4
[illegible]
ID   [illegible]     PRELIMINARY;     PRT;   [illegible] AA.
```

```
AC   [illegible]
DT   [illegible]
DT   [illegible]
DT   [illegible]
DE   [illegible]
OS   Homo sapiens (Human).
OC   [illegible]
OC   Mammalia; [illegible]
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=[illegible];
RA   [illegible]
RT   [illegible]
RL   [illegible]
DR   [illegible]
SQ   SEQUENCE [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy   [illegible] EVL [illegible]

Db   [illegible] EVL [illegible]


RESULT [illegible]
[illegible]
ID   [illegible]     PRELIMINARY;   [illegible] AA.
AC   [illegible]
DT   [illegible]
DT   [illegible]
DT   [illegible]
DE   [illegible]
RN   [illegible]
OS   Homo sapiens (Human).
OC   [illegible]
OC   Mammalia; [illegible]
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RA   [illegible]
RA   [illegible]
RA   [illegible]
RT   "[illegible]
RT   [illegible]"
RL   [illegible]
DR   EMBL; [illegible]
DR   MEROPS; [illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]
SQ   SEQUENCE [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy   [illegible]

Db   [illegible]


RESULT [illegible]
[illegible]
ID   [illegible]     PRELIMINARY;   [illegible] AA.
AC   [illegible]
DT   [illegible]
DT   [illegible]
DT   [illegible]
```

SUMMARIES

```
;        OPERATING SYSTEM: [illegible]
;        SOFTWARE: [illegible]
;      CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: [illegible]
;        FILING DATE: [illegible]
;      PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: [illegible]
;        FILING DATE: 20-APR-1994
;      PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: [illegible]
;        FILING DATE: [illegible]
;        CLASSIFICATION: [illegible]
;      ATTORNEY/AGENT INFORMATION:
;        NAME: [illegible], Stephanie
;        REGISTRATION NUMBER: [illegible]
;        REFERENCE/DOCKET NUMBER: [illegible]
;      TELECOMMUNICATION INFORMATION:
;        TELEPHONE: [illegible]
;        TELEFAX: [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]
;      SEQUENCE CHARACTERISTICS:
;        LENGTH: [illegible] amino acids
;        TYPE: amino acid
;        TOPOLOGY: linear
;      MOLECULE TYPE: protein
US-[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible]

Qy  [illegible] IVLGEPI [illegible]
Db  [illegible] EMIQEP [illegible]


RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;      APPLICANT: [illegible]
;      APPLICANT: [illegible], Chin Chun
;      TITLE OF INVENTION: HUMAN [illegible]
;      TITLE OF INVENTION: [illegible]
;      NUMBER OF SEQUENCES: [illegible]
;      CORRESPONDENCE ADDRESS:
;        ADDRESSEE: [illegible]
;        STREET: [illegible] Executive Square, [illegible]
;        CITY: [illegible]
;        STATE: CA
;        COUNTRY: U.S.A.
;        ZIP: [illegible]
;      COMPUTER READABLE FORM:
;        MEDIUM TYPE: [illegible]
;        COMPUTER: [illegible]
;        OPERATING SYSTEM: [illegible]
;        SOFTWARE: [illegible]
;      CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: [illegible]
;        FILING DATE: [illegible]-SEPT-[illegible]
;      PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: [illegible]
;        FILING DATE: [illegible]
;      PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: [illegible]
;        FILING DATE: [illegible]
;        CLASSIFICATION: [illegible]
;      ATTORNEY/AGENT INFORMATION:
;        NAME: [illegible], Stephanie
;        REGISTRATION NUMBER: [illegible]
;        REFERENCE/DOCKET NUMBER: [illegible]
;      TELECOMMUNICATION INFORMATION:
```

[illegible]

```
;       COMPUTER: IBM Compatible
;       OPERATING SYSTEM: MS [illegible]
;       SOFTWARE: Word Perfect [illegible]
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/[illegible]
;       FILING DATE: 26-FEB-[illegible]
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: JP 39563/1992
;       FILING DATE: 26-FEB-1992
;       APPLICATION NUMBER: JP [illegible]/1992
;       FILING DATE: 30-JUN-1992
;       APPLICATION NUMBER: JP [illegible]/1992
;       FILING DATE: 12-AUG-1992
;       APPLICATION NUMBER: JP [illegible]/1992
;       FILING DATE: 13-NOV-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Hamblet, C.Bruce
;       REGISTRATION NUMBER: [illegible]
;       REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (212) [illegible]
;       TELEFAX: (212) [illegible]
;   INFORMATION FOR SEQ ID NO: [illegible]:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 1464 amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single strand
;       TOPOLOGY: linear
;     MOLECULE TYPE: protein
;     ORIGINAL SOURCE:
;       ORGANISM: mouse
;       TISSUE TYPE: cerebellum
;     PUBLICATION INFORMATION:
;       AUTHORS: Masayoshi MISHINA
;       TITLE: NOVEL PROTEINS AND GENES CODING THE SAME
;       RELEVANT RESIDUES IN SEQ ID NO: [illegible] 1464
RQ [illegible] 26-[illegible]

  Query Match [illegible]  Score [illegible]  Length [illegible]
  Best Local Similarity [illegible]  Pred. No. [illegible]
  Matches [illegible]  Conservative [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

Qy        [illegible] EVIQEE [illegible]
             [illegible]
Db        [illegible] EMIQEE [illegible]


RESULT [illegible]
US-[illegible]A [illegible]
; Sequence [illegible], Application US/[illegible]A
; Patent No. [illegible]
;   GENERAL INFORMATION:
;     APPLICANT: [illegible]
;     APPLICANT: [illegible], Steven [illegible]
;     APPLICANT: [illegible], Chen W.
;     APPLICANT: [illegible]
;     TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR
;     TITLE OF INVENTION: SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND [illegible]
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Brown, Martin, Haller & McClain
;       STREET: [illegible] Union Street
;       CITY: San Diego
;       STATE: CA
;       COUNTRY: U.S.A.
;       ZIP: [illegible]
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/[illegible]A
```

```
[illegible]
;     ATTORNEY/AGENT INFORMATION:
;       NAME: [illegible]
[illegible]
;   INFORMATION FOR SEQ ID NO: [illegible]:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: [illegible]
;       TYPE: [illegible]
[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy        [illegible]

Db        [illegible]


RESULT [illegible]
US-[illegible]
; Sequence [illegible]
; Patent No. [illegible]
;   GENERAL INFORMATION:
;     APPLICANT: [illegible]
;     TITLE OF INVENTION: [illegible]
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: [illegible]
[illegible]
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: [illegible]
[illegible]
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]
[illegible]
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]
[illegible]
;     ATTORNEY/AGENT INFORMATION:
;       NAME: [illegible]
[illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: [illegible]
;       TELEFAX: [illegible]
;   INFORMATION FOR SEQ ID NO: [illegible]:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: [illegible]
[illegible]
```

[illegible]

[illegible]

RESULT [illegible]
US-08-[illegible]-7A-4
; Sequence 1, Application US/08[illegible]A
; Patent No. [illegible]
; GENERAL INFORMATION:
;   APPLICANT: [illegible]
;   APPLICANT: [illegible]
;   TITLE OF INVENTION: HUMAN [illegible]LIPASE [illegible] ALPHA AND DNA
;   TITLE OF INVENTION: SEQUENCE ENCODING THE SAME
;   NUMBER OF SEQUENCES: [illegible]
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: [illegible] FISH & NEAVE
;   STREET: [illegible] Avenue of the Americas
;   CITY: New York
;   STATE: New York
;   COUNTRY: [illegible]A
;   ZIP: [illegible]
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: [illegible]/MS-DOS
;   SOFTWARE: [illegible]
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/[illegible]7A
;   FILING DATE:
;   CLASSIFICATION: [illegible]
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: [illegible]
;   FILING DATE: 24-SEP-[illegible]
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PCT/JP9[illegible]
;   FILING DATE: 22-SEP-[illegible]
;   ATTORNEY/AGENT INFORMATION:
;   NAME: HALEY [illegible]
;   REGISTRATION NUMBER: [illegible]
;   REFERENCE/DOCKET NUMBER: [illegible]
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (212) [illegible]
;   TELEFAX: (212) [illegible]
; INFORMATION FOR SEQ ID NO: [illegible]:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: [illegible] amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-[illegible]A-[illegible]

Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

[illegible]
Db [illegible] EVMQEEF [illegible]

RESULT [illegible]
US-[illegible]A
; Sequence [illegible], Application US/[illegible]A
; Patent No. [illegible]
; GENERAL INFORMATION:
;   APPLICANT: [illegible]
;   APPLICANT: [illegible]
;   APPLICANT: [illegible]
;   APPLICANT: [illegible]
;   TITLE OF INVENTION: PLANT GLUTAMATE RECEPTORS
;   NUMBER OF SEQUENCES: [illegible]

[illegible]
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: [illegible]
;   ATTORNEY/AGENT INFORMATION:
;   NAME: [illegible]
;   REGISTRATION NUMBER: [illegible]
;   REFERENCE/DOCKET NUMBER: [illegible]
; INFORMATION FOR SEQ ID NO: [illegible]
;   MOLECULE TYPE: [illegible]

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible]

RESULT [illegible]
[illegible]
; GENERAL INFORMATION:
;   APPLICANT: [illegible]
;   APPLICANT: [illegible]
;   APPLICANT: [illegible]
;   APPLICANT: [illegible]
[illegible]
;   CURRENT APPLICATION DATA:
[illegible]
;   ATTORNEY/AGENT INFORMATION:
;   NAME: [illegible]
;   REGISTRATION NUMBER: [illegible]
;   REFERENCE/DOCKET NUMBER: [illegible]

SUMMARIES

ALIGNMENTS

[illegible]

TITLE OF INVENTION: [illegible] HUMAN [illegible] ALPHA AND DNA [illegible]

[illegible] SEQUENCE ENCODING THE SAME

[illegible]

CURRENT APPLICATION DATA:

[illegible]

PRIOR APPLICATION DATA:

[illegible]

PRIOR APPLICATION DATA:

[illegible]

ATTORNEY/AGENT INFORMATION:

NAME: [illegible]

REGISTRATION NUMBER: [illegible]

REFERENCE/DOCKET NUMBER: [illegible]

TELECOMMUNICATION INFORMATION:

[illegible]

INFORMATION FOR SEQ ID NO: [illegible]

SEQUENCE CHARACTERISTICS:

[illegible]

MOLECULE TYPE: [illegible]

[illegible]

```
;  APPLICANT: [illegible] Jean [illegible]
;  TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: US/[illegible]
;  CURRENT FILING DATE: 2001-10-12
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: PatentIn version [illegible]
; SEQ ID NO [illegible]
;   LENGTH: [illegible]
;   TYPE: [illegible]
;   ORGANISM: [illegible] sapiens
US-[illegible]

  Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
  Best Local Similarity [illegible]; Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible] VIQEEE [illegible]

Db [illegible] VIQEEE [illegible]


RESULT 12
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. US[illegible]A1
; GENERAL INFORMATION:
;  APPLICANT: Qi, Steve
;  APPLICANT: Akinsanya, Karen
;  APPLICANT: Ritter, Pierre
;  APPLICANT: [illegible], Jean Louis
;  TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: US/[illegible]
;  CURRENT FILING DATE: [illegible]-12
;  PRIOR APPLICATION NUMBER: US [illegible]
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: PatentIn version [illegible]
; SEQ ID NO [illegible]
;   LENGTH: 882
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-[illegible]

  Query Match [illegible]; Score [illegible]; DB [illegible]; Length 882;
  Best Local Similarity [illegible]; Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible] VIQEEE [illegible]
[illegible]
Db [illegible] VIQEEE [illegible]


RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. US[illegible]A1
; GENERAL INFORMATION:
;  APPLICANT: Qi, Steve
;  APPLICANT: Akinsanya, Karen
;  APPLICANT: Ritter, Pierre
;  APPLICANT: [illegible], Jean Louis
;  TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: US/[illegible]
;  CURRENT FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: PatentIn version [illegible]
```

```
[illegible]
US-[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy [illegible]

Db [illegible]


RESULT [illegible]
US-[illegible]
; Sequence [illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  TITLE OF INVENTION: [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: [illegible]
;  CURRENT FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: [illegible]
; SEQ ID NO [illegible]
;   LENGTH: [illegible]
;   TYPE: [illegible]
;   ORGANISM: [illegible]
US-[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy [illegible]

Db [illegible]


RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;  APPLICANT: Qi, Steve
;  APPLICANT: Akinsanya, [illegible]
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  TITLE OF INVENTION: NOVEL [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: [illegible]
;  CURRENT FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: [illegible]
; SEQ ID NO [illegible]
;   LENGTH: [illegible]
;   TYPE: [illegible]
;   ORGANISM: [illegible]
US-[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]
```

[illegible]

; Sequence [illegible], Application US/[illegible]6674
; Patent No. US20[illegible]15843A1
; GENERAL INFORMATION:
; APPLICANT: [illegible]
; APPLICANT: Akinsanya, Karen
; APPLICANT: [illegible], Pierre
; APPLICANT: [illegible]
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO [illegible]
; FILE REFERENCE: [illegible]
; CURRENT APPLICATION NUMBER: US/09/976,6[illegible]
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US [illegible]/24[illegible]
; PRIOR FILING DATE: 20[illegible]
; NUMBER OF SEQ ID NOS: [illegible]
; SOFTWARE: PatentIn version [illegible]
; SEQ ID NO [illegible]
; LENGTH: [illegible]
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-6[illegible]

[illegible]

RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. US[illegible]15843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: [illegible], Pierre
; APPLICANT: [illegible]
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO [illegible]
; FILE REFERENCE: [illegible]
; CURRENT APPLICATION NUMBER: US/[illegible]
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US [illegible]/24[illegible]
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: [illegible]
; SOFTWARE: PatentIn version [illegible]
; SEQ ID NO [illegible]
; LENGTH: [illegible]

[illegible]

OM protein - protein search, using sw model

Published Applications AA:*

SUMMARIES

Sequence 1, Appl

ALIGNMENTS

RESULT 1

; GENERAL INFORMATION

; APPLICANT:

; TITLE OF INVENTION:

; NUMBER OF SEQ ID NOS:

; SOFTWARE:

; LENGTH:

; TYPE:

; ORGANISM:

Query Match

Best Local Similarity

Matches

RESULT 2

; GENERAL INFORMATION

; APPLICANT:

; TITLE OF INVENTION:

>FD

```
;  APPLICANT: [illegible] Jean [illegible]
;  TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: US/[illegible]
;  CURRENT FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: PatentIn version [illegible]
; SEQ ID NO 26
;   LENGTH: 2411
;   TYPE: DNA
;   ORGANISM: Homo sapiens
[illegible]

Alignment Scores:
Pred. No.:            [illegible]     Length:        [illegible]
Score:                [illegible]     Matches:       [illegible]
Percent Similarity:   [illegible]     Conservative:  [illegible]
Best Local Similarity: 8[illegible].71%  Mismatches:  [illegible]
Query Match:          94.2[illegible]  Indels:       [illegible]
DB:                   1[illegible]     Gaps:          [illegible]

SW[illegible] (1[illegible]) x US[illegible] (1[illegible])

[illegible]
Db  [illegible] TATATAGAAAGA[illegible]

RESULT [illegible]
US[illegible]
; Sequence 4, Application US/[illegible]
; Patent No. US[illegible]
; GENERAL INFORMATION:
;  APPLICANT: [illegible]
;  APPLICANT: Akiyama, Karen
;  APPLICANT: [illegible], Pierre
;  APPLICANT: [illegible]
;  TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO [illegible]
;  FILE REFERENCE: 70669
;  CURRENT APPLICATION NUMBER: US/[illegible]/976,674
;  CURRENT FILING DATE: 2001-[illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: PatentIn version [illegible]
; SEQ ID NO 4
;   LENGTH: 2[illegible]
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US[illegible]976-674-4

Alignment Scores:
Pred. No.:            [illegible]     Length:        [illegible]
Score:                [illegible]     Matches:       [illegible]
Percent Similarity:   [illegible]     Conservative:  [illegible]
Best Local Similarity: [illegible]    Mismatches:    [illegible]
Query Match:          94.2[illegible]  Indels:       [illegible]
DB:                   1               Gaps:          [illegible]

SW464-4 (1[illegible]) x US[illegible]976-674-4 (1[illegible]2617)

[illegible]
Db  [illegible] TATATAGGAAAGA[illegible]

RESULT [illegible]
US[illegible]
; Sequence [illegible], Application US/[illegible]
; Publication No. US2002019767[illegible]A1
; GENERAL INFORMATION:
;  APPLICANT: [illegible] Y. [illegible]
;  APPLICANT: [illegible]

;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  TITLE OF INVENTION: [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: [illegible]
;  CURRENT FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: [illegible]
; SEQ ID NO [illegible]
;   LENGTH: [illegible]
;   TYPE: DNA
;   ORGANISM: [illegible]
[illegible]

Alignment Scores:
Pred. No.:            [illegible]
Score:                [illegible]
Percent Similarity:   [illegible]
Best Local Similarity: [illegible]
Query Match:          [illegible]
DB:                   [illegible]

SW[illegible]

[illegible]
Db  [illegible]

RESULT [illegible]
US[illegible]
; Sequence [illegible], Application [illegible]
; [illegible]
; GENERAL INFORMATION:
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  TITLE OF INVENTION: [illegible]
```

```
Percent Similarity:     100.0%        Conservative:    [illegible]
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            100.0%        Indels:          0
DB:                     1[illegible]  Gaps:            0

SWOPE464-1 (1-[illegible]) x US-09-[illegible]-624-12 (1-4[illegible])

Qy     1 [illegible]
         [illegible]
Db   [illegible] [illegible]AAGAAGAAT[illegible]

RESULT [illegible]
US-[illegible]-12
; Sequence 12, Application US/[illegible]
; Patent No. [illegible]A1
; GENERAL INFORMATION:
;  APPLICANT: [illegible], Steve
;  APPLICANT: Akinsanya, Karen
;  APPLICANT: [illegible], Pierre
;  APPLICANT: [illegible], Jean [illegible]
;  TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: US/[illegible]
;  CURRENT FILING DATE: 2001-1[illegible]
;  PRIOR APPLICATION NUMBER: US [illegible]/2[illegible],117
;  PRIOR FILING DATE: 2000-1[illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
;   LENGTH: 4[illegible]
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-[illegible]-12

Alignment Scores:
Pred. No.:              [illegible]   Length:          4[illegible]
Score:                  [illegible]   Matches:         [illegible]
Percent Similarity:     100.0%        Conservative:    [illegible]
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            100.0%        Indels:          0
DB:                     1[illegible]  Gaps:            0

SWOPE464-1 (1-[illegible]) x US-[illegible]-12 (1-[illegible])

Qy     1 [illegible]
         [illegible]
Db   [illegible] [illegible]AAGAAGAAT[illegible]

RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. US2002[illegible]A1
; GENERAL INFORMATION:
;  APPLICANT: [illegible] et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA[illegible]
;  CURRENT APPLICATION NUMBER: US/[illegible]
;  CURRENT FILING DATE: 2001-[illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]-12
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: PatentIn Ver. 2.[illegible]
; SEQ ID NO [illegible]
;   LENGTH: [illegible]
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: ([illegible])
;   OTHER INFORMATION: n equals a,t,c or g
;   NAME/KEY: misc_feature
;   LOCATION: ([illegible])
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: ([illegible])
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: ([illegible])
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: ([illegible])
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: ([illegible])
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-[illegible]

Alignment Scores:
Pred. No.:              [illegible]   Length:          [illegible]
Score:                  [illegible]   Matches:         [illegible]
Percent Similarity:     [illegible]   Conservative:    1
Best Local Similarity:  [illegible]   Mismatches:      0
Query Match:            [illegible]   Indels:          0
DB:                     [illegible]   Gaps:            0

SWOPE464-1 (1-[illegible]) x US-[illegible] (1-[illegible])

Qy     1 [illegible]
         [illegible]
Db   [illegible] [illegible]AGAAT[illegible]

RESULT 8
US-10-044-[illegible]
; Sequence 81[illegible], Application US/10044[illegible]
; Patent No. US[illegible]A1
; GENERAL INFORMATION:
;  APPLICANT: [illegible]
;  TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR [illegible] ACTIVATION
;  FILE REFERENCE: [illegible]US
;  CURRENT APPLICATION NUMBER: US/10/044,[illegible]
;  CURRENT FILING DATE: 2002-[illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: [illegible]
; SEQ ID NO 8[illegible]
;   LENGTH: [illegible]
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: [illegible]
;   OTHER INFORMATION: [illegible]
US-10-044-[illegible]

Alignment Scores:
Pred. No.:              [illegible]   Length:          [illegible]
Score:                  [illegible]   Matches:         [illegible]
Percent Similarity:     [illegible]   Conservative:    1
Best Local Similarity:  [illegible]   Mismatches:      [illegible]
Query Match:            [illegible]   Indels:          0
DB:                     [illegible]   Gaps:            0

SWOPE464-4 (1-[illegible]) x US-10-044-[illegible]-84[illegible] (1-[illegible])

Qy     1 [illegible]
         [illegible]
Db   [illegible] [illegible]AAGAAT[illegible]

RESULT 9
US-09-[illegible]
; Sequence [illegible], Application US/[illegible]624
; Patent No. US20[illegible]A1
; GENERAL INFORMATION:
;  APPLICANT: [illegible]
;  APPLICANT: Akinsanya, Karen
;  APPLICANT: [illegible], Pierre
```

```
Query Match [illegible] Indels [illegible]
DB: [illegible] Gaps: [illegible]

SW-[illegible]

[illegible]

RESULT 2
US-[illegible]
; Sequence 24, Application US/[illegible]
; Patent No. US[illegible]A1
; GENERAL INFORMATION:
;  APPLICANT: [illegible], Steve
;  APPLICANT: Akinsanya, Karen
;  APPLICANT: [illegible], Pierre
;  APPLICANT: [illegible]
;  TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: US/09/[illegible]
;  CURRENT FILING DATE: 2001-10-12
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: PatentIn version [illegible]
; SEQ ID NO [illegible]
;   LENGTH: [illegible]
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-[illegible]

Alignment Scores:
Pred. No.: [illegible] Length: [illegible]
Score: [illegible] Matches: [illegible]
Percent Similarity: [illegible] Conservative: [illegible]
Best Local Similarity: [illegible] Mismatches: [illegible]
Query Match: [illegible] Indels: [illegible]
DB: [illegible] Gaps: [illegible]

SW-[illegible]

Qy [illegible]
Db [illegible]AAGAAGAA[illegible]

RESULT 3
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. US[illegible]A1
; GENERAL INFORMATION:
;  APPLICANT: [illegible], Steve
;  APPLICANT: Akinsanya, Karen
;  APPLICANT: [illegible], Pierre
;  APPLICANT: Jourlen, Jean Louis
;  TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: [illegible]
;  CURRENT FILING DATE: 2001-10-12
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: PatentIn version [illegible]
; SEQ ID NO [illegible]
;   LENGTH: [illegible]
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-[illegible]

Alignment Scores:
Pred. No.: [illegible] Length: [illegible]
Score: [illegible] Matches: [illegible]
Percent Similarity: [illegible] Conservative: [illegible]

[illegible]
Query Match [illegible]
DB: [illegible]

SW-[illegible]

Qy [illegible]

Db [illegible]

RESULT 4
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;  APPLICANT: [illegible]
;  APPLICANT: Akinsanya, Karen
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  TITLE OF INVENTION: [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: [illegible]
;  CURRENT FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: [illegible]
; SEQ ID NO [illegible]
;   LENGTH: [illegible]
;   TYPE: DNA
;   ORGANISM: [illegible]
US-[illegible]

Alignment Scores:
Pred. No.: [illegible]
Score: [illegible]
Percent Similarity: [illegible]
Best Local Similarity: [illegible]
Query Match: [illegible]
DB: [illegible]

SW-[illegible]

Qy [illegible]

Db [illegible]

RESULT 5
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  TITLE OF INVENTION: [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: [illegible]
;  CURRENT FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: [illegible]
; SEQ ID NO [illegible]
;   LENGTH: [illegible]
;   TYPE: DNA
;   ORGANISM: [illegible]
US-[illegible]

Alignment Scores:
Pred. No.: [illegible]
Score: [illegible]
```

Minimum Match

Maximum Match

Command line parameters:

MODEL frame+ p2n.model DEV x11

DB Published Applications NA -QFMT=fastap -SUFFIX p2n.rnpb -MINMATCH

MATRIX

THR MAX

THR MIN MINLEN

MAXLEN

WARN

SUMMARIES

ALIGNMENTS

GENERAL

APPLICANT:

APPLICANT:

APPLICANT:

APPLICANT:

FILE REFERENCE:

CURRENT APPLICATION NUMBER:

CURRENT FILING DATE:

NUMBER OF SEQ ID NOS:

SOFTWARE:

LENGTH:

TYPE: DNA

ORGANISM:

Alignment Scores:

Matches:

Mismatches:

```
[illegible]
Best Local Similarity:  [illegible]        Mismatches: [illegible]
Query Match:            8[illegible].71%   Indels: [illegible]
DB:                     1                 Gaps: [illegible]

[illegible]

Qy      [illegible]
Db      [illegible]

RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application US/[illegible]A
; Patent No. [illegible]
;  GENERAL INFORMATION:
;    APPLICANT: [illegible]
;    APPLICANT: Ellis, Steven B.
;    APPLICANT: [illegible], [illegible] W.
;    APPLICANT: [illegible]
;    TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR
;    TITLE OF INVENTION: SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND [illegible]
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Brown, Martin, Haller & McClain
;      STREET: 1660 Union Street
;      CITY: San Diego
;      STATE: CA
;      COUNTRY: U.S.A.
;      ZIP: 921[illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.2[illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/[illegible]A
;      FILING DATE: [illegible]-APR-[illegible]
;      CLASSIFICATION: [illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: US [illegible]
;      FILING DATE: [illegible]
;      CLASSIFICATION: [illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Seidman, Stephanie
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: [illegible]
;      TELEFAX: [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 4[illegible] base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: both
;      TOPOLOGY: [illegible]
;    MOLECULE TYPE: [illegible]NA
;    FEATURE:
;      NAME/KEY: [illegible]
;      LOCATION: [illegible]
US-[illegible]

Alignment Scores:
[illegible]
Best Local Similarity:  71.43%             Mismatches: [illegible]
Query Match:            8[illegible].71%   Indels: [illegible]
DB:                     2                  Gaps: [illegible]

[illegible]

Qy      [illegible]
```

[illegible]

```
;           APPLICATION NUMBER: [illegible]
;           FILING DATE: 2[illegible] Apr [illegible]
;     PRIOR APPLICATION DATA:
;           APPLICATION NUMBER: [illegible]
;           FILING DATE: [illegible]
;     ATTORNEY/AGENT INFORMATION:
;           NAME: [illegible]
;           REGISTRATION NUMBER: [illegible]
;           REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;           TELEPHONE: 617-[illegible]
;           TELEFAX: 617 [illegible]
;  INFORMATION FOR SEQ ID NO: 4[illegible]:
;     SEQUENCE CHARACTERISTICS:
;           LENGTH: [illegible] base pairs
;           TYPE: nucleic acid
;           STRANDEDNESS: single
;           TOPOLOGY: linear
;     MOLECULE TYPE: DNA (genomic)
;     SEQUENCE DESCRIPTION: SEQ ID NO: 4[illegible]
US-[illegible]

Alignment Scores:
Pred. No.:                 2[illegible]       Length:          [illegible]
Score:                     31.00      Matches:         6
Percent Similarity:        100.0%     Conservative:    [illegible]
Best Local Similarity:     [illegible]     Mismatches:      
Query Match:               [illegible]     Indels:          
DB:                        1          Gaps:            

SW-[illegible] (1-[illegible]) x US-[illegible] (1-[illegible])

Qy          1 [illegible]
                [illegible]
Db         2[illegible] [illegible]AAAAAGA[illegible]

RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application US/0045[illegible]2B
; Patent No. [illegible]
;  GENERAL INFORMATION:
;     APPLICANT: Blattner, Frederick R.
;                Burland, Valerie
;                [illegible], N[illegible]
;                Plunkett, [illegible]
;                Welch, Rod
;     TITLE OF INVENTION: No. [illegible]
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;           ADDRESSEE: [illegible]
;           STREET: [illegible] South Pinckney Street
;           CITY: Madison
;           STATE: WI
;           COUNTRY: US
;           ZIP: [illegible]
;     COMPUTER READABLE FORM:
;           MEDIUM TYPE: Diskette, [illegible]
;           COMPUTER: IBM PC compatible
;           OPERATING SYSTEM: PC-DOS/MS-DOS
;           SOFTWARE: Word Perfect 8.0
;     CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/09/45[illegible]
;           FILING DATE: [illegible]
;           CLASSIFICATION: [illegible]
;     PRIOR APPLICATION DATA:
;           APPLICATION NUMBER: 60/110,[illegible]
;           FILING DATE: 04 DEC 1998
;     ATTORNEY/AGENT INFORMATION:
;           NAME: Seay, Nicholas [illegible]
;           REGISTRATION NUMBER: [illegible]
;           REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;           TELEPHONE: (608) [illegible]
```

```
[illegible]
;     SEQUENCE CHARACTERISTICS:
;           LENGTH: [illegible]
;           TYPE: [illegible]
;           STRANDEDNESS: [illegible]
;           TOPOLOGY: [illegible]
;     MOLECULE TYPE: [illegible] DNA
;     SEQUENCE DESCRIPTION: [illegible]
US-[illegible]

Alignment Scores:
[illegible]

SW-[illegible]

Qy [illegible]
Db [illegible]

RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application [illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;  APPLICANT: [illegible]
;  TITLE OF INVENTION: [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: [illegible]
;  CURRENT FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: [illegible]
; SEQ ID NO [illegible]
;   LENGTH: [illegible]
;   TYPE: DNA
;   ORGANISM: [illegible]
;   FEATURE:
;   OTHER INFORMATION: [illegible]
US-[illegible]

Alignment Scores:
[illegible]

SW-[illegible]

Qy [illegible]
Db [illegible]

RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application [illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;    APPLICANT: [illegible]
;    APPLICANT: [illegible]
;    APPLICANT: [illegible]
;    APPLICANT: [illegible]
;    TITLE OF INVENTION: [illegible]
;    TITLE OF INVENTION: [illegible]
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: [illegible]
```

```
;       TELECOMMUNICATION INFORMATION:
;         TELEPHONE:  (212) 596-9000
;         TELEFAX:  (212) 596-9090
;   INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;         LENGTH:  1933 base pairs
;         TYPE:  nucleic acid
;         STRANDEDNESS:  double
;         TOPOLOGY:  linear
;       MOLECULE TYPE:  cDNA to mRNA
;       FEATURE:
;         NAME/KEY:  CDS
;         LOCATION:  68..1585
US-08-627-907A-1

Alignment Scores:
Pred. No.:                  53.4          Length:        [illegible]
Score:                      33.00         Matches:       6
Percent Similarity:         100.00%       Conservative:  1
Best Local Similarity:      85.71%        Mismatches:    0
Query Match:                94.29%        Indels:        0
DB:                         3             Gaps:          0

SWOPE464-4 (1-7) x US-08-627-907A-1 (1-1933)

Qy          1 [illegible] 7
                [illegible]
Db       1073 [illegible]GTCATGCAGGAGGA[illegible] 1093

RESULT 2
US-08-627-907A-3
; Sequence 3, Application US/08627907A
; Patent No. 6060302
;   GENERAL INFORMATION:
;       APPLICANT:  HIRANO, Naoto
;       APPLICANT:  HIRAI, Hisamaru
;       TITLE OF INVENTION:  HUMAN PHOSPHOLIPASE C ALPHA AND DNA
;       TITLE OF INVENTION:  SEQUENCE ENCODING THE SAME
;       NUMBER OF SEQUENCES:  8
;       CORRESPONDENCE ADDRESS:
;         ADDRESSEE:  c/o FISH & NEAVE
;         STREET:  1251 Avenue of the Americas
;         CITY:  New York
;         STATE:  New York
;         COUNTRY:  USA
;         ZIP:  10020
;       COMPUTER READABLE FORM:
;         MEDIUM TYPE:  Floppy disk
;         COMPUTER:  IBM PC compatible
;         OPERATING SYSTEM:  PC-DOS/MS-DOS
;         SOFTWARE:  PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;         APPLICATION NUMBER:  US/08/627,907A
;         FILING DATE:
;         CLASSIFICATION:  435
;       PRIOR APPLICATION DATA:
;         APPLICATION NUMBER:  JP 5-238102
;         FILING DATE:  24-SEP-1993
;       PRIOR APPLICATION DATA:
;         APPLICATION NUMBER:  PCT/JP94/01572
;         FILING DATE:  22-SEP-1994
;       ATTORNEY/AGENT INFORMATION:
;         NAME:  HALEY Jr., James F.
;         REGISTRATION NUMBER:  27,794
;         REFERENCE/DOCKET NUMBER:  SHON 10
;       TELECOMMUNICATION INFORMATION:
;         TELEPHONE:  (212) 596-9000
;         TELEFAX:  (212) 596-9090
;   INFORMATION FOR SEQ ID NO: 3:
;       SEQUENCE CHARACTERISTICS:
;         LENGTH:  1942 base pairs
;         TYPE:  nucleic acid
;         STRANDEDNESS:  double
```

```
;         TOPOLOGY:  linear
;       MOLECULE TYPE:  DNA [illegible]
;       FEATURE:
;         NAME/KEY:  CDS
;         LOCATION:  [illegible]
US-08-[illegible]

Alignment Scores:
Pred. No.:                  [illegible]
Score:                      [illegible]   Matches:       [illegible]
Percent Similarity:         [illegible]
Best Local Similarity:      [illegible]
Query Match:                [illegible]
DB:                         [illegible]

SWOPE464-4 (1-[illegible]) x [illegible]

Qy          [illegible]
Db          [illegible]

RESULT 3
US-[illegible]
; Sequence [illegible], Application [illegible]
; Patent No. [illegible]
;   GENERAL INFORMATION:
;       APPLICANT:  [illegible]
;       TITLE OF INVENTION:  [illegible]
;       NUMBER OF SEQUENCES:  [illegible]
;       CORRESPONDENCE ADDRESS:
;         ADDRESSEE:  [illegible]
;         STREET:  [illegible]
;         CITY:  [illegible]
;         STATE:  CA
;         COUNTRY:  USA
;         ZIP:  [illegible]
;       COMPUTER READABLE FORM:
;         MEDIUM TYPE:  Diskette
;         COMPUTER:  IBM Compatible
;         OPERATING SYSTEM:  [illegible]
;         SOFTWARE:  [illegible]
;       CURRENT APPLICATION DATA:
;         APPLICATION NUMBER:  [illegible]
;         FILING DATE:  [illegible]
;         CLASSIFICATION:  [illegible]
;       PRIOR APPLICATION DATA:
;         APPLICATION NUMBER:  [illegible]
;         FILING DATE:  [illegible]
;       PRIOR APPLICATION DATA:
;         APPLICATION NUMBER:  [illegible]
;         FILING DATE:  [illegible]
;       PRIOR APPLICATION DATA:
;         APPLICATION NUMBER:  [illegible]
;         FILING DATE:  [illegible]
;       PRIOR APPLICATION DATA:
;         APPLICATION NUMBER:  [illegible]
;         FILING DATE:  [illegible]
;       ATTORNEY/AGENT INFORMATION:
;         NAME:  [illegible]
;         REGISTRATION NUMBER:  [illegible]
;         REFERENCE/DOCKET NUMBER:  [illegible]
;       TELECOMMUNICATION INFORMATION:
;         TELEPHONE:  [illegible]
;         TELEFAX:  [illegible]
;         TELEX:  [illegible]
;   INFORMATION FOR SEQ ID NO: [illegible]
;       SEQUENCE CHARACTERISTICS:
;         LENGTH:  [illegible] base pairs
;         TYPE:  nucleic acid
;         STRANDEDNESS:  [illegible]
;         TOPOLOGY:  [illegible]
;       MOLECULE TYPE:  DNA [illegible]
;       HYPOTHETICAL:  NO
```

[illegible]
Copyright (c) [illegible] [illegible] Ltd.

[illegible] search, using frame_plus_p2n model

Run on: February [illegible] (without alignments)
[illegible] Million cell updates/sec

Title: SWOPE464-4
Perfect score: [illegible]
Sequence: [illegible]

Scoring table: BLOSUM62
Xgapop [illegible], Xgapext [illegible]
Ygapop [illegible], Ygapext [illegible]
Fgapop [illegible], Fgapext [illegible]
Delop [illegible], Delext [illegible]

Searched: [illegible] seqs, [illegible] residues

Total number of hits satisfying chosen parameters: [illegible]

Minimum DB seq length: [illegible]
Maximum DB seq length: [illegible]

Post-processing: Minimum Match [illegible]%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
MODEL=frame+_p2n.model -DEV=xlh
[illegible]/SWOPE464/[illegible] [illegible]/app_query.fasta [illegible]
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=[illegible] [illegible]
[illegible] -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
[illegible] -ALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=[illegible]
-USER=SWOPE464 [illegible] -NCPU=[illegible] [illegible]
-NO_XLPY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=[illegible]
-WARN_TIMEOUT=[illegible] -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=[illegible] -FGAPEXT=[illegible]
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=[illegible] -DELEXT=[illegible]

Database: Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/[illegible]A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/[illegible]B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/[illegible]A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/[illegible]B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	[illegible]	[illegible]	[illegible]	[illegible]	US-08-627-[illegible]A-1	Sequence 1, Appl[illegible]
2	[illegible]	[illegible]	[illegible]	[illegible]	US-08-627-[illegible]A-[illegible]	Sequence [illegible], Appl[illegible]
3	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
4	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], Appl[illegible]
5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], Appl[illegible]
6	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], Appl[illegible]
7	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
8	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
9	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], Appl[illegible]
10	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
11	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], Appl[illegible]
12	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], Appl[illegible]
13	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
14	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
15	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
16	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
17	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
18	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], Appl[illegible]
19	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], Appl[illegible]
20	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], Appl[illegible]
21	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
22	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
23	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
24	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
25	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
26	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
27	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
28	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
29	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
30	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
31	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
32	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
33	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
34	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
35	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
36	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
37	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
38	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
39	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence [illegible], App[illegible]
40	[illegible]	[illegible]	[illegible]	[illegible]	US-08-[illegible]A-[illegible]	Sequence [illegible], App[illegible]
41	[illegible]	[illegible]	[illegible]	[illegible]	US-07-[illegible]	Sequence [illegible], App[illegible]
42	[illegible]	[illegible]	[illegible]	[illegible]	US-08-[illegible]	Sequence [illegible], App[illegible]
43	[illegible]	[illegible]	[illegible]	[illegible]	US-08-[illegible]	Sequence [illegible], App[illegible]
44	[illegible]	[illegible]	[illegible]	[illegible]	US-08-[illegible]	Sequence [illegible], App[illegible]
45	[illegible]	[illegible]	[illegible]	[illegible]	US-08-[illegible]	Sequence [illegible], App[illegible]

ALIGNMENTS

RESULT 1
US-08-627-[illegible]A-1
; Sequence 1, Application US/08627[illegible]A
; Patent No. [illegible]
; GENERAL INFORMATION:
; APPLICANT: [illegible]
; APPLICANT: [illegible]
; TITLE OF INVENTION: HUMAN PHOSPHOLIPASE [illegible] ALPHA AND DNA
; TITLE OF INVENTION: SEQUENCE ENCODING THE SAME
; NUMBER OF SEQUENCES: [illegible]
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: [illegible] & NEAVE
; STREET: [illegible] Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: [illegible]
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/627,[illegible]A
; FILING DATE:
; CLASSIFICATION: [illegible]
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP [illegible]
; FILING DATE: [illegible] SEP 199[illegible]
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP94/[illegible]
; FILING DATE: [illegible] 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY [illegible], James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: SHIN [illegible]

[illegible]

; [illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
; APPLICANT: [illegible]
; APPLICANT: Akinsanya, Karen
; APPLICANT: [illegible]
; APPLICANT: [illegible]
; TITLE OF INVENTION: NOVEL SERINE [illegible]
; FILE REFERENCE: [illegible]
; CURRENT APPLICATION NUMBER: [illegible]
; CURRENT FILING DATE: [illegible]
; PRIOR APPLICATION NUMBER: [illegible]
; PRIOR FILING DATE: [illegible]
; NUMBER OF SEQ ID NOS: [illegible]
; SOFTWARE: PatentIn version [illegible]
; SEQ ID NO [illegible]
; LENGTH: [illegible]
; TYPE: DNA
; ORGANISM: Homo sapiens

Alignment Scores:
Pred. No.: [illegible] Length: [illegible]
Score: [illegible] Matches: [illegible]
Percent Similarity: [illegible] Conservative: [illegible]
Best Local Similarity: 100.00% Mismatches: [illegible]
Query Match: [illegible] Indels: [illegible]
DB: [illegible] Gaps: [illegible]

[illegible]

; Sequence [illegible] Application [illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
; APPLICANT: [illegible]
; APPLICANT: Akinsanya, Karen
; APPLICANT: [illegible]
; APPLICANT: [illegible]
; TITLE OF INVENTION: NOVEL SERINE [illegible]
; FILE REFERENCE: [illegible]
; CURRENT APPLICATION NUMBER: [illegible]
; CURRENT FILING DATE: [illegible]
; PRIOR APPLICATION NUMBER: [illegible]
; PRIOR FILING DATE: [illegible]
; NUMBER OF SEQ ID NOS: [illegible]
; SOFTWARE: PatentIn version [illegible]
; SEQ ID NO [illegible]
; LENGTH: [illegible]
; TYPE: DNA
; ORGANISM: [illegible]

Alignment Scores:
[illegible]

SUMMARIES

7FD

PROV 1/18/00

ALIGNMENTS

GENERAL INFORMATION

7FD

7FD

[illegible]

```
;  APPLICANT: [illegible], David K.
;  APPLICANT: [illegible], Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED [illegible]
;  TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY [illegible] ARRAY
;  FILE REFERENCE: [illegible] X-1
;  CURRENT APPLICATION NUMBER: US/[illegible]/864,[illegible]
;  CURRENT FILING DATE: 2001-05-2[illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: US [illegible]
;  PRIOR FILING DATE: 20[illegible]
;  PRIOR APPLICATION NUMBER: US [illegible]
;  PRIOR FILING DATE: 200[illegible]
;  PRIOR APPLICATION NUMBER: GB 24[illegible]
;  PRIOR FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: 20[illegible]
;  PRIOR APPLICATION NUMBER: PCT/US01/[illegible]
;  PRIOR FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: PCT/US[illegible]
;  PRIOR FILING DATE: 2001-0[illegible]
;  PRIOR APPLICATION NUMBER: PCT/US[illegible]
;  PRIOR FILING DATE: 2001-0[illegible]
;  PRIOR APPLICATION NUMBER: PCT/US[illegible]
;  PRIOR FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: PCT/US[illegible]
;  PRIOR FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: PCT/US01/[illegible]
;  PRIOR FILING DATE: 2[illegible]
;  PRIOR APPLICATION NUMBER: PCT/U[illegible]
;  PRIOR FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: PCT/[illegible]
;  PRIOR FILING DATE: 2001-[illegible]
;  PRIOR APPLICATION NUMBER: PCT/US[illegible]
;  PRIOR FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]/[illegible]4[illegible]
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US [illegible]
;  PRIOR FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]/[illegible]4[illegible]
;  PRIOR FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: 4[illegible]17
;  SOFTWARE: [illegible] Sequence Listing Engine [illegible]
; SEQ ID NO [illegible]
;   LENGTH: [illegible]
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO [illegible]
;   OTHER INFORMATION: EXPRESSED IN [illegible]
;   OTHER INFORMATION: EXPRESSED IN FETAL [illegible]
;   OTHER INFORMATION: SWISSPROT [illegible] EVALUE [illegible]
;   OTHER INFORMATION: EST_HUMAN HIT: AA7[illegible] EVALUE 0,[illegible]
;   OTHER INFORMATION: EST_HUMAN HIT: AA[illegible] EVALUE [illegible]
[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]; Pred. No. [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]
```

[illegible]

[illegible]

```
GENERAL INFORMATION:
 APPLICANT: NAKAGAWA, [illegible]
 APPLICANT: [illegible]
 [illegible]
 APPLICANT: [illegible]
 [illegible]
 PRIOR APPLICATION NUMBER: [illegible]
 PRIOR FILING DATE: [illegible]
 [illegible]
 SOFTWARE: [illegible]
 [illegible]
 LENGTH: [illegible]
 TYPE: [illegible]
 [illegible]
```

[illegible]

```
; Sequence [illegible] Application US/[illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;  APPLICANT: [illegible]
;  [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: [illegible]
;  CURRENT FILING DATE: [illegible]
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: [illegible]
; SEQ ID NO [illegible]
;   LENGTH: [illegible]
;   TYPE: [illegible]
;   ORGANISM: [illegible]
;   FEATURE:
;   NAME/KEY: [illegible]
;   LOCATION: [illegible]
;   OTHER INFORMATION: Xaa [illegible] of the naturally occurring [illegible]
;   NAME/KEY: [illegible]
;   LOCATION: [illegible]
;   OTHER INFORMATION: Xaa [illegible] of the naturally occurring [illegible]
;   NAME/KEY: [illegible]
;   LOCATION: [illegible]
;   OTHER INFORMATION: Xaa [illegible] of the naturally occurring [illegible]
;   NAME/KEY: [illegible]
;   LOCATION: [illegible]
;   OTHER INFORMATION: Xaa [illegible] of the naturally occurring [illegible]
;   NAME/KEY: [illegible]
;   LOCATION: [illegible]
;   OTHER INFORMATION: Xaa [illegible] of the naturally occurring [illegible]
```

[illegible]

```
; GENERAL INFORMATION:
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
```

```
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: US/09/[illegible]
;  CURRENT FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: US [illegible]
;  PRIOR FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: Patentin version [illegible]
; SEQ ID NO [illegible]
;  LENGTH: [illegible]
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-09-976-674-[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy [illegible]

Db [illegible]


RESULT [illegible]
US-09-976-674-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. US[illegible]
; GENERAL INFORMATION:
;  APPLICANT: [illegible]
;  APPLICANT: Akiskalyan, Karen
;  APPLICANT: [illegible], Pierre
;  APPLICANT: [illegible], Jean-Louis
;  TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: US/09/976,674
;  CURRENT FILING DATE: 2001-10-12
;  PRIOR APPLICATION NUMBER: US [illegible]
;  PRIOR FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: Patentin version [illegible]
; SEQ ID NO [illegible]
;  LENGTH: [illegible]
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-09-976-674-[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy [illegible]

Db [illegible]


RESULT 4
US-09-976-674-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. US[illegible]
; GENERAL INFORMATION:
;  APPLICANT: [illegible]
;  APPLICANT: Akiskalyan, Karen
;  APPLICANT: [illegible], Pierre
;  APPLICANT: [illegible], Jean-Louis
;  TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: US/09/976,674
;  CURRENT FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: US [illegible]
;  PRIOR FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: Patentin version [illegible]
; SEQ ID NO [illegible]
;  LENGTH: [illegible]
```

[illegible]

Minimum Match

Maximum Match

Published Applications AA:*

ALIGNMENTS

RESULT 1

; Patent No.

; GENERAL INFORMATION:

; APPLICANT:

; APPLICANT:

; APPLICANT:

; APPLICANT:

; TITLE OF INVENTION:

; FILE REFERENCE:

; CURRENT APPLICATION NUMBER:

; CURRENT FILING DATE:

; PRIOR APPLICATION NUMBER:

; NUMBER OF SEQ ID NOS:

; SOFTWARE:

; SEQ ID NO

; LENGTH:

; TYPE:

RESULT 2

; Patent No.

; GENERAL INFORMATION:

; APPLICANT:

; APPLICANT:

; APPLICANT:

; APPLICANT:

; TITLE OF INVENTION:

```
;     APPLICANT: [illegible], Wilbert Frank [illegible]
;     APPLICANT: Bottaro, Wilhelm; [illegible], Lloyd [illegible]
;     TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
;     TITLE OF INVENTION: ALPHA, AND USES THEREOF
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Felfe & Lynch
;       STREET: [illegible] Third Avenue
;       CITY: New York City
;       STATE: New York
;       COUNTRY: USA
;       ZIP: [illegible]
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Diskette, 3.5 inch, [illegible] Mb storage
;       COMPUTER: IBM PS/2
;       OPERATING SYSTEM: [illegible]
;       SOFTWARE: Wordperfect
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: [illegible]-MARCH-[illegible]
;       CLASSIFICATION: [illegible]
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]/2[illegible]
;       FILING DATE: 20-APRIL-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Hanson, Norman D.
;       REGISTRATION NUMBER: [illegible]
;       REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (212) [illegible]
;       TELEFAX: (212) [illegible]
;   INFORMATION FOR SEQ ID NO: [illegible]:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: [illegible] amino acids
;       TYPE: amino acid
;       TOPOLOGY: linear
[illegible]289A-2

[illegible]
[illegible]
[illegible]

[illegible] ENV[illegible]
[illegible]
[illegible] CONVERT [illegible]


RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. 5[illegible]
;   GENERAL INFORMATION:
;     APPLICANT: [illegible], Wilbert Frank [illegible]
;     APPLICANT: Bottaro, Wilhelm; [illegible]
;     TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION
;     TITLE OF INVENTION: PROTEIN ALPHA, AND USES THEREOF
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Felfe & Lynch
;       STREET: [illegible] Third Avenue
;       CITY: New York City
;       STATE: New York
;       COUNTRY: USA
;       ZIP: [illegible]
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Diskette, [illegible] Mb storage
;       COMPUTER: IBM PS/2
;       OPERATING SYSTEM: [illegible]
;       SOFTWARE: Wordperfect
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/[illegible]
;       FILING DATE: [illegible] 1997
;       CLASSIFICATION: [illegible]
```

```
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: [illegible]
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: [illegible]
;     ATTORNEY/AGENT INFORMATION:
;       NAME: [illegible]
;       REGISTRATION NUMBER: [illegible]
;       REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: [illegible]
;       TELEFAX: [illegible]
;   INFORMATION FOR SEQ ID NO: [illegible]:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: [illegible]
;       TYPE: [illegible]
;       TOPOLOGY: [illegible]
[illegible]

[illegible]
[illegible]
[illegible]

[illegible] ENV[illegible]
[illegible]
[illegible] CONVERT [illegible]

[illegible]
[illegible]
```

```
;    OTHER INFORMATION: Sequence
US-08-[illegible]-12

  Query Match            64.[illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity  [illegible];  Pred. No. [illegible];
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy    [illegible] NVHE[illegible]
Db    [illegible] MEESKHEIH [illegible]44


RESULT [illegible]
US-[illegible]-[illegible]
; Sequence 2[illegible], Application US/[illegible]
; Patent No. [illegible]44
;  GENERAL INFORMATION:
;       APPLICANT: BRUCKDORFER, KARL R
;                  [illegible], [illegible]
;       TITLE OF INVENTION: ANTICOAGULANT [illegible] FRAGMENTS DERIVED
;                           FROM APOLIPOPROTEIN [illegible]
;       NUMBER OF SEQUENCES: [illegible]
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: NIXON & VANDERHYE [illegible]
;            STREET: [illegible] North Glebe R[illegible]
;            CITY: ARLINGTON
;            STATE: VA
;            COUNTRY: USA
;            ZIP: 22[illegible]
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.[illegible]
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/[illegible]
;            FILING DATE: 07-[illegible]-1998
;            CLASSIFICATION: [illegible]
;       ATTORNEY/AGENT INFORMATION:
;            NAME: SADOFF, B.J.
;            REGISTRATION NUMBER: [illegible]
;            REFERENCE/DOCKET NUMBER: [illegible]
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: [illegible]
;            TELEFAX: 70[illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 4536 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: [illegible]:
US-[illegible]-[illegible]

  Query Match            64.7%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity  [illegible];  Pred. No. [illegible];
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy    [illegible] NVHEAH [illegible]
Db    [illegible] NKE[illegible] 1412


RESULT [illegible]
US-[illegible]-2
; Sequence 2, Application US/08[illegible]
; Patent No. [illegible]2[illegible]
;  GENERAL INFORMATION:
;    APPLICANT: Willnow, [illegible]
;    APPLICANT: [illegible]
;    APPLICANT: [illegible]
;    APPLICANT: Mueller, [illegible]
```

[illegible]

```
  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy      [illegible] NVHFAH [illegible]

Db      [illegible] NLHFAH [illegible]


RESULT [illegible]
US-[illegible]-182-728A-[illegible]
; Sequence [illegible], Application US/[illegible]182728A
; Patent No. [illegible]
;  GENERAL INFORMATION:
;    APPLICANT: Wang, Bruce
;    APPLICANT: Fisher, Joseph
;    APPLICANT: [illegible]
;    TITLE OF INVENTION: [illegible]
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Flehr, Hohbach, Test, Albritton
;      ADDRESSEE: & Herbert
;      STREET: Four Embarcadero Center, [illegible]
;      CITY: San Francisco
;      STATE: California
;      COUNTRY: United States
;      ZIP: [illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #[illegible], Version #[illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/[illegible]/182,728A
;      FILING DATE: 07-JUN-1995
;      CLASSIFICATION: 435
;    ATTORNEY/AGENT INFORMATION:
;      NAME: [illegible], Robin M.
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: A-[illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (415) [illegible]
;      TELEFAX: (415) [illegible]
;      TELEX: [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: [illegible] amino acids
;      TYPE: amino acid
;      STRANDEDNESS: unknown
;      TOPOLOGY: unknown
;    MOLECULE TYPE: protein
US-[illegible]-182-728A-[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy      [illegible] ENVHFAH [illegible]

Db      [illegible] ENFHYKH [illegible]


RESULT [illegible]
US-[illegible]-A-[illegible]
; Sequence [illegible], Application US/[illegible]A
; Patent No. [illegible]
;  GENERAL INFORMATION:
;    APPLICANT: [illegible]
;    APPLICANT: [illegible]
;    TITLE OF INVENTION: [illegible]
;    TITLE OF INVENTION: [illegible]
;    NUMBER OF SEQUENCES: [illegible]
```

```
RP   STRAIN [illegible];
RX   MEDLINE [illegible]; PubMed [illegible];
RA   [illegible] H., Nakamura Y., [illegible] S., Kanek[illegible], [illegible];
RA   [illegible] S.;
RT   "Structural analysis of Arabidopsis [illegible]
RT   sequence features of the regions of [illegible]
RT   physically assigned BAC clones.";
RL   DNA Res. 4:[illegible](1997).
DR   EMBL; AB[illegible]; BAB[illegible]; -.
DR   InterPro; IPR[illegible]; DEFV_N [illegible].
DR   InterPro; IPR[illegible]; Peptidase [illegible].
DR   InterPro; IPR[illegible]; Ser [illegible].
DR   Pfam; PF[illegible]; DEFV_N_[illegible].
DR   Pfam; PF[illegible]; Peptidase [illegible].
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible] CRC64;

  Query Match [illegible];
  Best Local Similarity   77.8%;  Pred. No. [illegible];
  Matches    7;  Conservative    1;  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy   [illegible] NVHAH [illegible]
Db   [illegible] NVH[illegible]


RESULT [illegible]
[illegible]
ID   [illegible]   PRELIMINARY;   PRT;   [illegible] AA.
AC   [illegible];
DT   [illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   Hypothetical [illegible] kDa protein.
GN   [illegible].
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN [illegible] N2;
RX   MEDLINE [illegible]; PubMed [illegible];
RA   Waterston R.;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN BRISTOL N2;
RA   [illegible] S., [illegible] S.;
RT   "The sequence of C. elegans [illegible].";
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN BRISTOL N2;
RA   Waterston R.;
RL   Submitted (A[illegible]-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; [illegible]; AAA[illegible]; -.
KW   Hypothetical protein.
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible] CRC64;

  Query Match [illegible];
  Best Local Similarity   77.8%;  Pred. No. [illegible];
  Matches    7;  Conservative    1;  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy   [illegible] HENVHHAH [illegible]
Db   [illegible] HENVHHAH [illegible]4


Search completed: [illegible]
Job time : [illegible] secs
```

```
[illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  [illegible]

     [illegible] ENVHEAH [illegible]

     [illegible] ENVHEH [illegible]

[illegible]
ID   [illegible]      PRELIMINARY;      PRT;   [illegible] AA.
AC   [illegible];
DT   [illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   Hypothetical [illegible] protein (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RA   Strausberg R.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; [illegible]; AAH[illegible].1; -.
DR   [illegible]
DR   InterPro; [illegible]; Peptidase_S[illegible].
DR   InterPro; [illegible]; Ser_[illegible]_site.
DR   Pfam; PF[illegible]; Peptidase_S[illegible]; 1.
KW   Hypothetical protein.
FT   NON_TER      [illegible]      1
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  [illegible]

     [illegible] ENVHEAH [illegible]

     [illegible] ENVHEH [illegible]

[illegible]
ID   [illegible]      PRELIMINARY;      PRT;   [illegible] AA.
AC   [illegible];
DT   [illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   [illegible] (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; [illegible]
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RA   [illegible]
RL   Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; [illegible]; AA[illegible]
[illegible]
```

[illegible]

RA Nature [illegible]

RT "Genome sequence of the nematode C. elegans: a platform for [illegible]"

[illegible]

RA Waterston R. [illegible]

RL Submitted [illegible] to the EMBL/GenBank/DDBJ databases.

[illegible]

Query Match [illegible]
Best Local Similarity [illegible] Pred. No. [illegible]
Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

[illegible]

RA Kimura K., [illegible] Matsuo K., Nakamura Y., Sekine M.,
RA Kikuchi H., [illegible] Wanatsuma M., Murakawa K., Kanehori K.,
RA Takahashi [illegible] Shiba A., Sugiyama A., Kawakami B., [illegible]
RA [illegible] Masuda Y., Nagai K., [illegible]
RL Submitted [illegible] to the EMBL/GenBank/DDBJ databases.

[illegible]

```
ID   [illegible]   PRELIMINARY;   [illegible] AA.
DE   [illegible] peptidase [illegible] (Fragment).
OS   [illegible] sapiens (Human).
OC   Eukaryota; Metazoa; [illegible] Vertebrata; [illegible]
OC   Mammalia; Eutheria; Primates; [illegible] Homo.
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RX   MEDLINE [illegible] PubMed [illegible]
RT   "[illegible] expression and chromosomal [illegible]
RT   dipeptidyl peptidase (DPP) IV homolog, [illegible]";
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match   [illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy   [illegible] DENVHEAH [illegible]
Db   [illegible] DENVHEAH [illegible]


ID   [illegible]   PRELIMINARY;   [illegible] AA.
DT   [illegible] (TrEMBLrel. [illegible] Created)
DT   [illegible] (TrEMBLrel. [illegible] Last sequence update)
DT   [illegible] (TrEMBLrel. [illegible] Last annotation update)
DE   peptidase IV (EC [illegible]).
OS   Homo sapiens (Human).
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RA   [illegible] Hayashi K., [illegible]
RA   [illegible] Sugano S., [illegible]
RA   Watanabe M., [illegible] Kaku Y., [illegible]
RA   Takahashi M., Chiba Y., [illegible]
RA   Watanabe S., [illegible] Murakami K., [illegible]
RA   [illegible] Wakamatsu A., [illegible]
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (MAY-[illegible]) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK0[illegible]; BAB[illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match   [illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy   [illegible] DENVHEAH [illegible]
Db   [illegible] DENVHEAH [illegible]


[illegible]
ID   [illegible]
DE   [illegible]
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; [illegible] Vertebrata; [illegible]
RP   SEQUENCE FROM N.A.
SQ   SEQUENCE   [illegible] AA;

  Query Match   [illegible]

Qy   [illegible] DENVHEAH [illegible]
Db   [illegible]


[illegible]
ID   [illegible]
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; [illegible] Vertebrata; [illegible]
RP   SEQUENCE FROM N.A.
RC   STRAIN=[illegible]
RA   [illegible]
RA   Sakai K., [illegible]
RA   [illegible]
```

```
DR   EMBL; M22[illegible]; AAA412[illegible].1; [illegible].
DR   EMBL; AF[illegible]; AAG30[illegible].1; [illegible].
DR   PIR; A410[illegible]; A41[illegible].
DR   PIR; A43775; A43775.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 2.
DR   SMART; SM00248; ANK; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 1.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hydrolase; Transit peptide; Mitochondrion; ANK repeat.
FT   TRANSIT       1     16       MITOCHONDRION.
FT   CHAIN        17    674       GLUTAMINASE, KIDNEY ISOFORM 68 KDA.
FT   CHAIN        73    674       GLUTAMINASE, KIDNEY ISOFORM 65 KDA.
FT   REPEAT      [illegible]    619       ANK.
FT   DOMAIN       [illegible]     [illegible]       POLY-GLN.
FT   CONFLICT     28     28       T -> A (IN REF. [illegible]).
FT   CONFLICT    [illegible]49    351       [illegible]K -> EEG (IN REF. 2).
SQ   SEQUENCE   674 AA;  7402[illegible] MW;  4B2524A4[illegible] CRC64;

  Query Match             66.7%;  Score [illegible];  DB [illegible];  Length 674;
  Best Local Similarity   55.6%;  Pred. No. [illegible];
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    [illegible];  Gaps    [illegible];

Qy          1 [illegible]ENVHEAH 9
              [illegible]
Db        629 MDEALHEGH 637


Search completed: February [illegible], 2003, 12:[illegible]
Job time : 13.5806 secs
```

```
RC   TISSUE=Brain;
RA   Turner A., Wang [illegible], Hawkinson [illegible], Porter [illegible];
RT   "Identification and expression of human renal and hepatic glutaminase
RT   isoforms.";
RL   Submitted (JAN [illegible]) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. ([illegible]).
RC   TISSUE=Brain;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Holcomb T., Taylor L., Trohkimoinen [illegible], Curthoys N.P.;
RT   "Isolation, characterization and expression of a human brain
RT   mitochondrial glutaminase cDNA.";
RL   Brain Res. Mol. Brain Res. [illegible](2000).
RN   [illegible]
RP   SEQUENCE OF [illegible] FROM N.A.
RC   TISSUE=Colorectal carcinoma;
RA   Turner A., McGivan J.D.;
RT   "Adenoma and carcinoma cell lines derived from colorectal tumours
RT   express different isoforms of glutaminase.";
RL   Submitted (JUN 2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: CATALYZES THE FIRST REACTION IN THE PRIMARY PATHWAY
CC       FOR THE RENAL CATABOLISM OF GLUTAMINE.
CC   -!- CATALYTIC ACTIVITY: L-glutamine + H(2)O = L-glutamate + NH(3).
CC   -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; KGA (SHOWN HERE), GAM AND
CC       GAC; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: KGA IS EXPRESSED PREDOMINANTLY IN BRAIN AND
CC       KIDNEY BUT NOT IN LIVER. GAC IS EXPRESSED PRINCIPALLY IN CARDIAC
CC       MUSCLE AND PANCREAS BUT NOT IN LIVER OR BRAIN. AND GAM IS
CC       EXPRESSED SOLELY IN CARDIAC AND SKELETAL MUSCLE.
CC   -!- SIMILARITY: BELONGS TO THE GLUTAMINASE FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 ANK REPEAT.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; AB[illegible]; AAD[illegible].1; -.
DR   EMBL; AF[illegible]; AAF[illegible].1; -.
DR   EMBL; AF[illegible]; AAF[illegible].1; -.
DR   EMBL; AF[illegible]; AAF[illegible].1; -.
DR   EMBL; AB0[illegible]; BAA[illegible].1; -.
DR   EMBL; AF[illegible]; AAF[illegible].1; -.
DR   EMBL; AF[illegible]; AAD[illegible].1; -.
DR   EMBL; AF[illegible]; AAD[illegible].1; -.
DR   Genew; HGNC:[illegible]; GLS.
DR   MIM; [illegible];
DR   InterPro; IPR[illegible]; ANK.
DR   Pfam; PF00023; ank; 1.
DR   SMART; SM00248; ANK; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 1.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hydrolase; Transit peptide; Mitochondrion; ANK repeat;
KW   Alternative splicing.
FT   TRANSIT       1     [illegible]       MITOCHONDRION (BY SIMILARITY).
FT   CHAIN        [illegible]   [illegible]       GLUTAMINASE, KIDNEY ISOFORM.
FT   REPEAT       [illegible]   [illegible]       ANK.
FT   VARSPLICE    [illegible]   [illegible]       [illegible] (IN [illegible] FORM GAM).
FT   VARSPLICE    [illegible]   [illegible]       MISSING (IN ISOFORM GAM).
FT   VARSPLICE    [illegible]   [illegible]       VKSVINL[illegible]
FT                                    VAAAL[illegible]
FT                                    [illegible]
FT                                    [illegible]
FT                                    [illegible] (IN ISOFORM GAC).
FT   CONFLICT     [illegible]   [illegible]       [illegible] (IN REF. [illegible]; AAF[illegible]).
FT   CONFLICT     [illegible]   [illegible]       [illegible] (IN REF. [illegible]; AAF[illegible]).
FT   CONFLICT     [illegible]   [illegible]       [illegible] (IN REF. [illegible]).
FT   CONFLICT     [illegible]   [illegible]       V -> A (IN REF. [illegible]).
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy        [illegible]IENVHESR[illegible]
              [illegible]
Db      [illegible]MEFAIHE[illegible]


RESULT [illegible]
GLSK_RAT
ID   GLSK_RAT       STANDARD;      PRT;   [illegible] AA.
AC   P13264;
DT   01-JAN-[illegible] (Rel. [illegible], Created)
DT   01-DEC-[illegible] (Rel. [illegible], Last sequence update)
DT   16-OCT-2001 (Rel. [illegible], Last annotation update)
DE   Glutaminase, kidney isoform, mitochondrial precursor (EC 3.5.1.2)
DE   (GLS) (L-glutamine amidohydrolase) (K-glutaminase).
GN   GLS.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE OF [illegible].
RC   STRAIN=Sprague-Dawley; TISSUE=Kidney;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Shapiro R.A., [illegible], Srinivasan M., Curthoys N.P.;
RT   "Isolation, characterization, and in vitro expression of a cDNA that
RT   encodes the kidney isoenzyme of the mitochondrial glutaminase.";
RL   J. Biol. Chem. [illegible](1991).
RN   [2]
RP   SEQUENCE OF [illegible] FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Banner C., [illegible], Shapiro R.A., Wenthold R.J., Nakatani Y.,
RA   Lampel K.A., [illegible] W., Hale D., Curthoys N.P.;
RT   "Isolation of a cDNA for rat brain glutaminase.";
RL   Brain Res. [illegible](1988).
RN   [3]
RP   SEQUENCE OF [illegible] FROM N.A.
RC   STRAIN=Sprague-Dawley;
RA   Taylor L., [illegible], Newsome W., Shapiro R.A., Srinivasan M.,
RA   Curthoys N.P.;
RT   "Isolation and characterization of the promoter region of the rat
RT   kidney-type glutaminase gene.";
RL   Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   PROCESSING.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Perera S.Y., [illegible] M., Curthoys N.P.;
RT   "Biosynthesis and processing of mitochondrial glutaminase in HTC
RT   hepatoma cells.";
RL   Biochem. [illegible](19[illegible]).
CC   -!- FUNCTION: CATALYZES THE FIRST REACTION IN THE PRIMARY PATHWAY
CC       FOR THE RENAL CATABOLISM OF GLUTAMINE.
CC   -!- CATALYTIC ACTIVITY: L-glutamine + H(2)O = L-glutamate + NH(3).
CC   -!- SUBUNIT: [illegible] COMPOSED OF [illegible] AND [illegible] KDA PEPTIDES IN A [illegible]
CC       RATIO.
CC   -!- SUBCELLULAR LOCATION: Mitochondrial.
CC   -!- TISSUE SPECIFICITY: KIDNEY, BRAIN, AND INTESTINE.
CC   -!- SIMILARITY: BELONGS TO THE GLUTAMINASE FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 ANK REPEAT.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; M[illegible]; AAA[illegible].1; -.
```

```
MENI_E[illegible]
ID   MEND_[illegible]     STANDARD;      PRT;   [illegible] AA.
AC   P[illegible]; P[illegible]; [illegible];
DT   [illegible] (Rel. [illegible], Created)
DT   [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] (Rel. 41, Last annotation update)
DE   Menaquinone biosynthesis protein menD [illegible]: 2-succinyl-[illegible]
DE   2,4-cyclohexadiene-1-carboxylate synthase (EC [illegible]) (SHCHC
DE   synthase); 2-oxoglutarate decarboxylase (EC 4.[illegible]) (Alpha
DE   ketoglutarate decarboxylase) (KDC)].
GN   MEND [illegible] B2264.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible];
RT   "Sequence and overexpression of the menD gene from Escherichia coli."
RL   J. Bacteriol. [illegible](1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=9742[illegible]; PubMed=[illegible];
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., [illegible]
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Yamamoto Y., Aiba H., Baba T., Hayashi K., [illegible], Isono K.,
RA   Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA   Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA   Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasundaram S.,
RA   Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA   Yamagata S., Horiuchi T.;
RT   "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT   K-12 genome corresponding to [illegible] on the linkage map and
RT   analysis of its sequence features.";
RL   DNA Res. 4:91-113(1997).
RN   [4]
RP   SEQUENCE [illegible] FROM N.A.
RC   STRAIN=K12;
RA   Sharma V., Hudspeth M.E.S., Meganathan R.;
RL   Submitted (JUL-1994) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   CHARACTERIZATION.
RC   STRAIN=K12;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible] Sharma V., Hudspeth M.E.S., Meganathan R.;
RT   "Menaquinone (vitamin K2) biosynthesis: evidence that the Escherichia
RT   coli menD gene encodes both 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-
RT   carboxylic acid synthase and alpha-ketoglutarate decarboxylase
RT   activities.";
RL   J. Bacteriol. [illegible](199[illegible]).
CC   -!- CATALYTIC ACTIVITY: Isochorismate + 2-oxoglutarate =
CC       2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate + pyruvate
CC       + CO(2).
CC   -!- CATALYTIC ACTIVITY: 2-oxoglutarate = [illegible]
CC       + CO(2).
CC   -!- COFACTOR: Binds [illegible] thiamine pyrophosphate and [illegible]
CC       [illegible].
CC   -!- PATHWAY: Menaquinone biosynthesis; first step.
CC   -!- INDUCTION: BY ISOPENTYL-BETA-D-THIOGALACTOPYRANOSIDE (IPTG).
CC   -!- SIMILARITY: BELONGS TO THE TPP ENZYME FAMILY.
CC   -!- CAUTION: REF.[illegible] SEQUENCE DIFFERS [illegible] SHOWN IN [illegible]
CC       AND IS [illegible] (5[illegible] AA) [illegible] FRAMESHIFT.
CC   [illegible]
CC   [illegible]
CC   between the Swiss [illegible]
CC   [illegible]
CC   [illegible] see http://www.[illegible]
CC   or send an email to [illegible]
CC   [illegible]
DR   EMBL; M[illegible]; AAA[illegible]
DR   EMBL; AE[illegible]
DR   EMBL; [illegible]
DR   EMBL; [illegible]
DR   EMBL; [illegible]
DR   EMBL; [illegible]
DR   EMBL; [illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]
DR   Pfam; [illegible]
DR   [illegible]
KW   [illegible]
KW   [illegible]
FT   [illegible]
FT   [illegible]
FT   [illegible]
FT   [illegible]
SQ   SEQUENCE [illegible]

Matches [illegible]

[illegible]

RESULT [illegible]
[illegible]
ID   [illegible]
AC   [illegible]
DT   [illegible]
DT   [illegible]
DT   [illegible]
DE   [illegible]
DE   [illegible]
GN   [illegible]
OS   [illegible]
OC   [illegible] Metazoa; [illegible] Vertebrata; [illegible]
OC   Mammalia; Eutheria; [illegible]
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A. [illegible]
RC   TISSUE=[illegible];
RX   MEDLINE=[illegible]
RA   [illegible]
RT   [illegible]
RT   [illegible]
RL   [illegible]
RN   [2]
RP   SEQUENCE FROM N.A. [illegible]
RC   [illegible]
RX   MEDLINE=[illegible]
RA   [illegible]
RA   [illegible]
RT   "[illegible]
RT   [illegible]
RT   [illegible]
RL   [illegible]
RN   [3]
RP   SEQUENCE FROM N.A. [illegible]
```

```
KW   ...
DR   EMBL; ...
DR   MIM; ...
RT   "Cyclophilin 40: a protein with homology to the P59 component of the
RT   steroid receptor complex. Cloning of the cDNA and further
CC   -!- CATALYTIC ACTIVITY: Peptidylproline (omega=180) = peptidylpro...
CC   -!- ENZYME REGULATION: LESS SENSITIVE TO INHIBITION BY CYCLOSPORIN A
CC   -!- SIMILARITY: BELONGS TO THE CYCLOPHILIN-TYPE PPIASE FAMILY.
CC   -!- SIMILARITY: CONTAINS 3 TPR REPEATS.
CC   -!- SIMILARITY: CONTAINS 3 TPR REPEATS.
```

```
CC       [illegible] IN THE [illegible] CATALYTIC [illegible].
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein (Probable).
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; [illegible] (SHOWN HERE) AND [illegible]
CC       ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN BRAIN.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY [illegible].
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced [illegible]
CC   between the Swiss Institute of Bioinformatics and the EMBL [illegible]
CC   the European Bioinformatics Institute. [illegible]
CC   use by non-profit institutions as long as its content [illegible]
CC   modified and this statement is not removed. [illegible]
CC   entities requires a license agreement (See http://www.isb-sib.ch/[illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M[illegible]; AAA[illegible]; -.
DR   EMBL; M[illegible]; AAA[illegible]; -.
DR   MEROPS; [illegible]; -.
DR   Genew; HGNC:[illegible]; [illegible].
DR   MIM; [illegible]; -.
DR   InterPro; [illegible]; [illegible] N-term.
DR   InterPro; [illegible]; Peptidase [illegible].
DR   InterPro; [illegible]; Ser [illegible] site.
DR   Pfam; PF[illegible]; Peptidase [illegible]; 1.
DR   Pfam; PF[illegible]; [illegible]_N-term; 1.
KW   Transmembrane; Glycoprotein; Signal-anchor; Alternative splicing.
FT   DOMAIN        [illegible]   [illegible]   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      [illegible]   [illegible]   SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
FT                                            (POTENTIAL).
FT   DOMAIN        [illegible]   [illegible]   EXTRACELLULAR (POTENTIAL).
FT   CARBOHYD      [illegible]   [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD      [illegible]   [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD      [illegible]   [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD      [illegible]   [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD      [illegible]   [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD      [illegible]   [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD      [illegible]   [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   VARSPLIC      1             [illegible]   MASIY[illegible]
FT                                            [illegible]RA[illegible]AAA[illegible]
FT                                            [illegible]MTAKE[illegible] (IN ISOFORM [illegible]
FT                                            S).
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match            [illegible];  Score [illegible];
  Best Local Similarity  [illegible];  Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy         [illegible]NVH[illegible]
                        :
Db         [illegible]EKIHE[illegible]


RESULT [illegible]
[illegible]_HUMAN
ID   [illegible]_HUMAN     STANDARD;      PRT;   [illegible] AA.
AC   [illegible];
DT   [illegible] (Rel. [illegible], Created)
DT   [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] (Rel. [illegible], Last annotation update)
DE   [illegible] protein [illegible] ([illegible]) [illegible]
GN   [illegible] OR [illegible].
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible] M.;
RA   [illegible] N., [illegible] W.;
RT   "[illegible]
RT   [illegible]";
```

[illegible]

```
CC
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; M[illegible]; AAC[illegible]; -.
DR   EMBL; M[illegible]; AAC[illegible]; -.
DR   MEROPS; S09.[illegible]; -.
DR   InterPro; IPR[illegible]; DPPIV_N_term.
DR   InterPro; IPR[illegible]; Peptidase_S9.
DR   InterPro; IPR[illegible]; Ser_estrs_site.
DR   Pfam; PF00326; Peptidase_S9; 1.
DR   Pfam; PF00930; DPPIV_N_term; 1.
KW   Transmembrane; Glycoprotein; Signal-anchor; Alternative splicing.
FT   DOMAIN        1     [illegible]    CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    [illegible]   [illegible]    SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                              (POTENTIAL).
FT   DOMAIN      [illegible]   [illegible]    EXTRACELLULAR (POTENTIAL).
FT   CARBOHYD    [illegible]   [illegible]    N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    [illegible]   [illegible]    N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    [illegible]   [illegible]    N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    [illegible]   [illegible]    N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    [illegible]   [illegible]    N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    [illegible]   [illegible]    N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    [illegible]   [illegible]    N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    [illegible]   [illegible]    N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC      1     [illegible]    MASLYQRF[illegible]
FT                              [illegible]
FT                              AKEPNASGKSV[illegible] (IN ISOFORM DPPX-S).
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match            [illegible];  Score [illegible];  DB 1;  Length [illegible];
  Best Local Similarity  [illegible];  Pred. No. [illegible];
  Matches [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy          [illegible]ENVHFAH[illegible]
               [illegible]
Db      [illegible] EEKIHFQH[illegible]


RESULT [illegible]
DPPX_BOVIN
ID   DPPX_BOVIN     STANDARD;      PRT;   [illegible] AA.
AC   P42785;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Dipeptidyl peptidase IV like protein (Dipeptidyl aminopeptidase
DE   related protein) (Dipeptidylpeptidase VI) (DPPX).
GN   DPP6.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC   TISSUE=Brain;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Wada K., Yokotani N., Hunter C., Doi K., Wenthold R.J., Shimasaki S.;
RT   "Differential expression of two distinct forms of mRNA encoding
RT   members of a dipeptidyl aminopeptidase family.";
RL   Proc. Natl. Acad. Sci. U.S.A. [illegible]
CC   -!- FUNCTION: MAY BE INVOLVED IN THE PHYSIOLOGICAL PROCESSES OF BRAIN
CC       FUNCTION. HAS NO DIPEPTIDYL AMINOPEPTIDASE ACTIVITY. THE LACK OF
CC       ACTIVITY MAY BE DUE TO THE SUBSTITUTION OF AN ASPARTATE RESIDUE
CC       FOR THE SERINE RESIDUE IN THE PROPOSED CATALYTIC TRIAD.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC   -!- ALTERNATIVE PRODUCTS: [illegible] mRNAs DPPX-L (SHOWN HERE) AND DPPX-S;
CC       ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE BRAIN. DPPX-S
CC       IS EXPRESSED [illegible] IN THE BRAIN WHEREAS DPPX-S IS FOUND IN
CC       BRAIN, KIDNEY, OVARY AND TESTIS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S9B.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; M[illegible]; AAA[illegible]; -.
DR   EMBL; M[illegible]; AAA[illegible]; -.
DR   MEROPS; S09.[illegible]; -.
DR   InterPro; IPR[illegible]; DPPIV_N_term.
DR   InterPro; IPR[illegible]; Peptidase_S9.
DR   InterPro; IPR[illegible]; Ser_estrs_site.
DR   Pfam; PF00326; Peptidase_S9; 1.
DR   Pfam; PF00930; DPPIV_N_term; 1.
KW   Transmembrane; Glycoprotein; Signal-anchor; Alternative splicing.
FT   DOMAIN        1     [illegible]    CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    [illegible]   [illegible]    SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT                              (POTENTIAL).
FT   DOMAIN      [illegible]   [illegible]    EXTRACELLULAR (POTENTIAL).
FT   CARBOHYD    [illegible]   [illegible]    N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    [illegible]   [illegible]    N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    [illegible]   [illegible]    N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    [illegible]   [illegible]    N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    [illegible]   [illegible]    N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    [illegible]   [illegible]    N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC      1     [illegible]    MASLYQRFTGKINTSRSFPAPPEASHLLGGQ[illegible]
FT                              PLGAQAPAAAPRERGGGGAGGRPRFQYQARSD[illegible]
FT                              MTTAKEPNASGKSVQQQ[illegible] (IN ISOFORM DPPX-S).
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]2B84[illegible] CRC64;

  Query Match            [illegible];  Score [illegible];  DB 1;  Length [illegible];
  Best Local Similarity  [illegible];  Pred. No. [illegible];
  Matches [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy          [illegible]ENV[illegible]
               [illegible]
Db      [illegible] EKIH[illegible]


RESULT [illegible]
DPPX_HUMAN
ID   DPPX_HUMAN     STANDARD;      PRT;   [illegible] AA.
AC   P42658;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Dipeptidyl peptidase IV like protein (Dipeptidyl aminopeptidase
DE   related protein) (Dipeptidylpeptidase VI) (DPPX).
GN   DPP6.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Hippocampus;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Yokotani N., Doi K., Wenthold R.J., Wada K.;
RT   "Non-conservation of a catalytic residue in a dipeptidyl
RT   aminopeptidase IV-related protein encoded by a gene on human
RT   chromosome 7.";
RL   Hum. Mol. Genet. [illegible](1993).
CC   -!- FUNCTION: MAY BE INVOLVED IN THE PHYSIOLOGICAL PROCESSES OF BRAIN
CC       FUNCTION. HAS NO DIPEPTIDYL AMINOPEPTIDASE ACTIVITY. THE LACK OF
CC       ACTIVITY MAY BE DUE TO THE SUBSTITUTION OF AN ASPARTATE RESIDUE
```

```
RT   [illegible] "XPR4." [illegible]
RL   [illegible]
CC   -!- FUNCTION: [illegible]
CC       [illegible] INCREASING [illegible]
CC       INVOLVED IN [illegible] LYMPHOCYTES [illegible] BONE MARROW [illegible]
CC       CARDIAC VENTRICULAR SEPTUM FORMATION. PLAYS ALSO AN ESSENTIAL [illegible]
CC       IN VASCULARIZATION OF THE GASTROINTESTINAL TRACT [illegible]
CC       REGULATING VASCULAR BRANCHING AND/OR REMODELING PROCESSES IN
CC       ENDOTHELIAL CELLS. INVOLVED IN CEREBELLAR NEURONAL LAYER
CC       FORMATION, PREVENTING PREMATURE MIGRATION [illegible]
CC       CELLS FROM THE EXTERNAL GRANULE LAYER [illegible]
CC       MEDIATE HIPPOCAMPAL NEURON SURVIVAL.
CC   -!- SUBCELLULAR LOCATION: [illegible]
CC   -!- ALTERNATIVE PRODUCTS: [illegible] CXCR4-A [illegible]
CC       B/LESTR [illegible] (SHOWN HERE) ARE PRODUCED BY ALTERNATIVE [illegible]
CC   -!- TISSUE SPECIFICITY: LYMPHOCYTES, MACROPHAGES, NEUTROPHILS,
CC       MICROGLIAL CELLS AND ASTROCYTES. FOUND IN SPLEEN, THYMUS,
CC       BONE MARROW, LYMPH NODES AND, AT LOWER LEVELS, IN [illegible]
CC       INTESTINE, STOMACH AND KIDNEY. [illegible] PREDOMINANT IN ALL
CC       TISSUES TESTED.
CC   -!- DEVELOPMENTAL STAGE: HIGH EXPRESSION [illegible] EMBRYONIC DEVELOPMENT
CC       [illegible] SEEM TO BE ASSOCIATED WITH [illegible] ANY
CC       PARTICULAR CELL TYPE, BUT IS WIDELY [illegible] WHEN THERE IS A
CC       REQUIREMENT FOR CELL MOVEMENT. FREQUENTLY ASSOCIATED WITH [illegible]
CC       DIFFERENTIATED CELL TYPES AND DOWN REGULATED WITH SUBSEQUENT
CC       DIFFERENTIATION. DETECTED IN SITES WITH HAEMOPOIETIC POTENTIAL:
CC       THE YOLK SAC (7.5, [illegible] AND 12.5 DPC) AND FETAL LIVER (12.5 [illegible])
CC       DURING GASTRULATION, AT 7.5 [illegible] EXPRESSED IN THE [illegible]
CC       AND THE DEFINITIVE ENDODERM, AS GASTRULATION [illegible]
CC       [illegible]). EXPRESSION IN THE MESODERM IS DOWN REGULATED WHILE [illegible]
CC       BECOMES PREDOMINANT IN NEURAL ECTODERM. ENDODERMAL EXPRESSION [illegible]
CC       RETAINED IN THE [illegible] AND LATER IN A SUBSET OF [illegible]
CC       DERIVATIVES, INCLUDING THE STOMACH ([illegible] DPC), THE [illegible]
CC       THE GALLBLADDER AND THE LUNG EPITHELIUM (12.5 DPC). IN NEURONAL
CC       TISSUES: AT 10.5 AND 12.5 DPC, EXPRESSED IN THE DORSAL ROOT
CC       GANGLIA, IN THE VENTRAL MANTLE LAYER OF THE SPINAL CORD [illegible]
CC       [illegible]). IN THE HINDBRAIN, AT 14.5 DPC, EXPRESSION [illegible]
CC       [illegible] THE NEURAL EPITHELIUM [illegible] THE VENTRICULAR [illegible]
CC       [illegible] THE EXTERNAL GRANULAR LAYER OF THE VENTRAL [illegible]
CC       [illegible] CEREBELLUM). EXPRESSED IN [illegible]
CC       [illegible] AT 10.5 DPC AND IN THE DEVELOPING THALAMUS [illegible]
CC       [illegible] A LESSER EXTENT, THE DEVELOPING HYPOTHALAMUS. AT 14.5 DPC
CC       RESTRICTED TO THE REGION WHERE THALAMUS AND HYPOTHALAMUS MEET
CC       [illegible] IN A DISCRETE BAND OF CELLS AT THE EDGE OF THE [illegible]
CC       [illegible]. IN THE VASCULAR SYSTEM: EXPRESSED IN THE ENDOTHELIUM [illegible]
CC       NUMBER OF [illegible] VESSELS, BUT NOT ALL. AT [illegible] AND 12.5 [illegible]
CC       SUCH AS VITELLINE/ [illegible] VESSELS [illegible] VENTRICULAR WALL
CC       [illegible], FACIAL VESSELS AND AT [illegible], IN THE VASCULA[illegible]
CC       OF THE [illegible]. THE EXPRESSION SEEMS TO BE ASSOCIATED WITH
CC       EXPANDING VASCULAR NETWORKS. IN THE HEART DEVELOPMENT, EXPRESS[illegible]
CC       AT 10.5 DPC IN THE PRECURSOR TO THE A[illegible]PULMONARY (AP) SEPT[illegible]
CC       12.5 DPC, DETECTED IN THE AP SEPTUM A[illegible] THE BASE OF THE [illegible]
CC       [illegible] AND IN THE ATRIOVENTRICULAR VALVES. DETECTED IN [illegible]
CC       [illegible] FROM [illegible] DPC AT [illegible] AND [illegible] DPC, EXPRESS[illegible]
CC       THE RATHKE'S POUCH.
CC   -!- PTM: [illegible] (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY [illegible]
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement [illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]; AAB[illegible].1; -.
DR   EMBL; [illegible]; AAC[illegible].1; -.
DR   EMBL; [illegible]; BAA[illegible].1; -.
DR   EMBL; [illegible]; CAB[illegible].1; -.
DR   EMBL; [illegible]; CAB[illegible].1; -.
DR   EMBL; X[illegible]; CAA[illegible].1; -.
DR   EMBL; X[illegible]; CAA[illegible].1; -.
```

```
DR   EMBL; [illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]
KW   [illegible]
KW   [illegible]
FT   DOMAIN [illegible]
FT   TRANSMEM [illegible]
FT   DOMAIN [illegible]
FT   TRANSMEM [illegible]
FT   DOMAIN [illegible]
FT   TRANSMEM [illegible]
FT   DOMAIN [illegible]
FT   TRANSMEM [illegible]
FT   DOMAIN [illegible]
FT   TRANSMEM [illegible]
FT   DOMAIN [illegible]
FT   TRANSMEM [illegible]
FT   DOMAIN [illegible]
FT   TRANSMEM [illegible]
FT   DOMAIN [illegible]
FT   DISULFID [illegible]
FT   MOD_RES [illegible]
FT   CARBOHYD [illegible]
FT   VARSPLIC [illegible]
FT   [illegible]
SQ   SEQUENCE [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

[illegible]

[illegible]

RESULT [illegible]
[illegible]
ID   [illegible]
AC   [illegible]
DT   [illegible]
DT   [illegible]
DT   [illegible]
DE   [illegible]
GN   [illegible]
OS   [illegible]
OC   [illegible]
OC   [illegible]
OX   [illegible]
RN   [illegible]
RP   [illegible]
RC   [illegible]
RX   [illegible]
RA   [illegible]
RT   "Differential expression [illegible]
RT   [illegible]
RL   [illegible]
CC   -!- [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   -!- ALTERNATIVE PRODUCTS: [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
```

[illegible]

[illegible]

```
GN   Alt- [illegible] SALT[illegible].
OS   Vaccinia virus (strain WR).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
OX   NCBI_TaxID=[illegible]254;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 91[illegible]; PubMed 187[illegible];
RA   Blasco R., Cole N.B., Moss B.;
RT   "Sequence analysis, expression, and deletion of a vaccinia virus gene
RT   encoding a homolog of profilin, a eukaryotic actin-binding protein.";
RL   J. Virol. 65:4598-4608(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 9125[illegible]; PubMed 20[illegible];
RA   Smith G.L., Chan Y.S., Howard S.T.;
RT   "Nucleotide sequence of 42 kbp of vaccinia virus strain WR from near
RT   the right inverted terminal repeat.";
RL   J. Gen. Virol. 72:1349-1376(1991).
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; M724[illegible]; AAA48[illegible].1; -.
DR   EMBL; D11[illegible]; BAA01816.1; -.
DR   PIR; [illegible]; [illegible].
DR   PIR; JQ1780; JQ1780.
SQ   SEQUENCE   [illegible]4 AA;  22577 MW;  [illegible] CRC64;

  Query Match             68.6%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   66.7%;  Pred. No. [illegible];
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 LLENVHPAH 9
              [illegible]
Db         73 LLENVIPTH 81


RESULT [illegible]
GIDB_STRPY
ID   GIDB_STRPY     STANDARD;      PRT;   2[illegible] AA.
AC   Q9A[illegible];
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Methyltransferase gidB (EC 2.1.-.-) (Glucose inhibited division
DE   protein B).
GN   GIDB OR SPY[illegible].
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX   MEDLINE 211[illegible]684; PubMed 112[illegible];
RA   Ferretti J.J., McShan W.M., Ajdic D., Savic D.J., Savic G., Lyon K.,
RA   Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.,
RA   Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J.,
RA   Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT   "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
CC   -!- FUNCTION: Probable S-adenosyl-L-methionine-dependent
CC       methyltransferase specific for a sterol and/or lipid substrate (By
CC       similarity).
CC   -!- SIMILARITY: BELONGS TO THE GIDB FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. [illegible]
CC   [illegible]
DR   EMBL; [illegible]
DR   InterPro; [illegible]
DR   InterPro; [illegible]
DR   Pfam; [illegible]
DR   ProDom; [illegible]
DR   TIGRFAMs; [illegible]
KW   [illegible]
SQ   SEQUENCE   [illegible] AA; [illegible] MW; [illegible] CRC64;

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches    [illegible]

Qy          [illegible] NVHPAH [illegible]

Db          [illegible] NV[illegible]


RESULT [illegible]
[illegible]
ID   [illegible]
AC   [illegible]
DT   [illegible]
DT   [illegible]
DT   [illegible]
DE   [illegible]
GN   [illegible]
OS   [illegible]
OC   [illegible]
OC   [illegible]
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=[illegible]
RX   MEDLINE [illegible]
RA   [illegible]
RA   [illegible]
RA   [illegible]
RT   "[illegible]
RL   [illegible]
CC   -!- SIMILARITY: [illegible]
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. [illegible]
CC   between the Swiss Institute of Bioinformatics and the EMBL [illegible]
CC   the European Bioinformatics Institute. [illegible]
CC   [illegible]
CC   [illegible]
CC   entities requires a license agreement (See http://www.[illegible]
CC   or send an email to [illegible]).
CC   -----------------------------------------------------------------------
DR   EMBL; [illegible]
DR   InterPro; [illegible]
DR   Pfam; [illegible]
DR   TIGRFAMs; [illegible]
DR   PROSITE; [illegible]
DR   PROSITE; [illegible]
KW   [illegible]
SQ   SEQUENCE   [illegible] AA; [illegible] MW; [illegible] CRC64;

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches    [illegible]

Qy          1 LLENVHPAH [illegible]

Db          [illegible]
```

[illegible]

SUMMARIES

[illegible]

ALIGNMENTS

[illegible]

VAIV_VA[illegible]

[illegible]

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ [illegible]

[illegible]

VAIV_VA[illegible]

[illegible]

```
A;Description: The complete genomic sequence of the Modified Vaccinia [illegible] MVA) strai
A;Reference number: Z20877
A;Accession: T37428
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-19[illegible] ANT
A;Cross-references: EMBL:U94848; PIDN:AAB965[illegible]
A;Experimental source: strain Ankara
C;Genetics:
A;Note: MVA155R

  Query Match             68.6%;  Score 35;  DB 2;  Length 19[illegible];
  Best Local Similarity   66.7%;  Pred. No. [illegible];
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    [illegible];

Qy          1 LLENVHEAE 9
                [illegible]
Db         73 LLDNVFEIE 81


RESULT 1[illegible]
I42521
A45R protein - vaccinia virus (strain Copenhagen)
C;Species: vaccinia virus
A;Note: host Homo sapiens (man)
C;Date: [illegible] Nov-199[illegible] #sequence_revision [illegible] Nov 199[illegible] #text_change 08 Apr [illegible]
C;Accession: I42521
R;Johnson, G.P.
submitted to GenBank, June 199[illegible]
A;Reference number: A3[illegible]172
A;Accession: I42521
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-194 <JOH>

  Query Match             68.6%;  Score 35;  DB 2;  Length 194;
  Best Local Similarity   66.7%;  Pred. No. [illegible];
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    [illegible];

Qy          1 LLENVHEAE 9
                [illegible]
Db         73 LLDNVFEIE 81


RESULT 1[illegible]
JQ178[illegible]
SalF5R protein precursor - vaccinia virus (strain WR)
C;Species: vaccinia virus
C;Date: [illegible] Sep 199[illegible] #sequence_revision [illegible] Sep 199[illegible] #text_change 20[illegible]
C;Accession: JQ178[illegible]; [illegible]4[illegible]8[illegible]7
R;Smith, G.L.; Chan, Y.S.; Howard, S.T.
J. Gen. Virol. 72, 1349-1376, 1991
A;Title: Nucleotide sequence of 42kbp of vaccinia virus strain WR from near the right [illegible]
A;Reference number: JQ1767; MUID:91259063; PMID:[illegible]
A;Accession: JQ178[illegible]
A;Molecule type: DNA
A;Residues: 1-194 <SMI>
A;Cross-references: DDBJ:D11079; NID:g222717; PIDN:BAA01816.1; PID:[illegible]
R;Blasco, R.; Cole, N.B.; Moss, B.
J. Virol. 65, 45[illegible]-4608, 1991
A;Title: Sequence analysis, expression, and deletion of a vaccinia virus [illegible]
A;Reference number: A46897; MUID:91332995; PMID:[illegible]
A;Accession: [illegible]4[illegible]8[illegible]7
A;Molecule type: DNA
A;Residues: 1-194 <BLA>
A;Cross-references: GB:M72474; NID:g335761; PIDN:AAA48[illegible]
C;Keywords: glycoprotein
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-194/Product: SalF5R protein #status predicted <MAT>
F;65,114/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             68.6%;  Score 35;  DB 2;  Length 194;
  Best Local Similarity   66.7%;  Pred. No. [illegible];
```

```
  Matches    [illegible]

Qy          1 L[illegible]ENVHEAE 9
                [illegible]
Db         7[illegible] [illegible]DNVFEIE 8[illegible]


RESULT 1[illegible]
E82[illegible]
conserved hypothetical protein [illegible]
C;Species: Vibrio cholerae
C;Date: [illegible] Aug 200[illegible] #sequence_revision [illegible] #text_change [illegible]
C;Accession: E82[illegible]
R;Heidelberg, J.F.; Eisen, J.A.; [illegible]
chardson, D.; Ermolaeva, M.D.; [illegible]
[illegible] McKalines [illegible] Venter, [illegible] M.
Nature 406, [illegible]
A;Title: DNA sequence [illegible]
A;Reference number: A82[illegible] MUI[illegible]
A;Accession: E82[illegible]
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-[illegible] <HEI>
A;Cross-references: GB:AE[illegible]
A;Experimental source: [illegible]
C;Genetics:
A;Gene: VC[illegible]
A;Map position: [illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches    [illegible]

Qy          [illegible]ENVHEA[illegible]
                [illegible]
Db          [illegible]

Search completed: [illegible]
Job time : [illegible]
```

[illegible]

RESULT [illegible]

[illegible]

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

[illegible]

RESULT [illegible]

[illegible]

A;Molecule type: [illegible]NA

[illegible]

C;Superfamily: dipeptidyl peptidase IV

Query Match [illegible] Score [illegible] DB 2; Length [illegible]
Best Local Similarity [illegible] Pred. No. [illegible]
Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

[illegible]

RESULT 1[illegible]
[illegible]
pyruvate carboxylase [illegible] pycA [similarity] - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change [illegible]
C;Accession: [illegible]
R;Kunst, F.; [illegible] N.; Moszer, I.; Albertini, A.M.; Alloni, [illegible]
[illegible] Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, [illegible]
Nature [illegible]
A;Authors: [illegible] Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; [illegible]
[illegible] Hilbert, H.; Holsappel, S.; [illegible]
Koetter, P.; [illegible] G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, [illegible]
A;Authors: [illegible] V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; [illegible]
[illegible] A.; Oudega, B.; Park, S.H.; Parro, V.; [illegible]
Rieger, M.; [illegible] Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; [illegible]
A;Authors: [illegible] Schroeter, R.; Scoffone, F.; Sekiguchi, [illegible]
akeuchi, M.; [illegible] A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; [illegible]
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A;Authors: Yoshikawa, H.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
A;Reference number: [illegible] MUID:98044033; PMID:9384377
A;Accession: [illegible]
A;Status: preliminary; [illegible] acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: [illegible] GB:AL009126; NID:[illegible] PIDN:[illegible]
A;Experimental source: strain 168
C;Genetics:
A;Gene: pycA
C;Superfamily: [illegible] carboxylase; biotin carboxylase homology; [illegible]
C;Keywords: [illegible]
F;[illegible]/Domain: biotin carboxylase homology <BCH>
F;[illegible]/Domain: [illegible] /biotin binding homology <BIB>
F;[illegible]/Binding site: [illegible] (Lys) (covalent) #status predicted

Query Match [illegible] Score [illegible] DB 2; Length [illegible]
Best Local Similarity [illegible] Pred. No. [illegible]
Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

[illegible]

RESULT 1[illegible]
[illegible]
membrane [illegible] (strain Ankara)
C;Species: [illegible]
A;Variety: strain Ankara
C;Date: [illegible] #text_change [illegible]
C;Accession: [illegible]
R;Antoine, [illegible] Falkner, F.G.; [illegible]
submitted [illegible] Library, March 1997

```
Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy [illegible] ENVHAH [illegible]
Db [illegible] ENVHA [illegible]

RESULT [illegible]
[illegible]
hypothetical protein RC0605 [imported] - Rickettsia conorii (strain Malish 7)
C;Species: Rickettsia conorii
C;Date: [illegible] Sep-2001 #sequence_revision [illegible] Sep-2001 #text_change [illegible]
C;Accession: [illegible]
R;Ogata, H.; Audic, S.; Renesto-Audiffren, [illegible]
Science [illegible]
A;Title: Mechanisms of evolution in Rickettsia [illegible] and Rickettsia [illegible]
A;Reference number: [illegible]
A;Accession: [illegible]
A;Status: preliminary
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: [illegible] PID:AAL [illegible]
[illegible]
A;[illegible]

Query Match [illegible]
Best Local Similarity [illegible] Pred. No. [illegible]
Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy [illegible] DENVHEAH [illegible]
Db [illegible] DERVHEAY [illegible]

RESULT [illegible]
[illegible]
dipeptidyl aminopeptidase [illegible] fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: [illegible] #sequence_revision [illegible] #text_change [illegible]
C;Accession: T41763
R;Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August [illegible]
A;Reference number: [illegible]
A;Accession: [illegible]
A;Status: preliminary; translated from GB/EMBL [illegible]
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: EMBL:AL [illegible] PIDN:CAA [illegible]
A;Experimental source: strain [illegible] cosmid [illegible]
C;Genetics:
A;Gene: SPAC [illegible]
A;Map position: [illegible]
C;Superfamily: dipeptidyl-peptidase IV

Query Match [illegible]
Best Local Similarity [illegible] Pred. No. [illegible]
Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy [illegible] ENVHEAH [illegible]
Db [illegible] ENVHE [illegible]

RESULT [illegible]
[illegible]
hypothetical protein [illegible] Synechocystis sp. (strain [illegible])
C;Species: Synechocystis sp.
A;Variety: [illegible]
C;Date: [illegible]-Apr-[illegible] #sequence_revision [illegible]-Apr-[illegible] #text_change [illegible]
C;Accession: S76[illegible]
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; [illegible]
[illegible]
DNA Res. [illegible]
```

[illegible]

XX
KW antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;
KW Grave's disease; neurological disorder; ...
XX
OS Homo sapiens.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Meyers RA, Williamson M;
XX
PT New polypeptides 21956, member of human prolyl ... peptidase ...
PT useful as diagnostic targets and therapeutic agents for ...
PT cancers ... and leukemia ...
XX
CC This invention relates to an isolated ... human prolyl
CC ... peptidase, which is cytostatic, antidiabetic, antiarthritic,
CC neuroprotective, antithyroid, dermatological, antipsoriatic,
CC antiasthmatic, ophthalmological, antiinflammatory, nootropic,
CC antiparkinsonian, anticonvulsant, ... vasotropic,
CC antiinflammatory, cardiant, antiatherosclerotic ... and
CC metabolic in its action. Uses include gene therapy, expression ...
CC activity of 21956 protein modulator. It is useful for identifying ...
CC compound which binds to it and can be used in preventing, treating
CC or detecting a cellular proliferative or differentiative disorder.
CC The 21956 molecules can act as novel diagnostic targets and therapeutic
CC agents for controlling disorders associated with the aberrant ...
CC ... peptidase ... associated with ...
CC differentiation and proliferation ... function ...
CC reproductive, neurological and cardiovascular function. The ...
CC ... are thus useful for treating and preventing ...
CC proliferative and differentiative disorders ...
CC disorders, immune disorders such as autoimmune diseases, diabetes
CC mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,
CC neuronal disorders, ... diseases, vascular disorders ...
CC metabolic ... disorders. This sequence represents the ...
CC acid sequence of 21956 human prolyl ... peptidase.
XX
SQ Sequence ... AA;

Query Match ... ; Score ... ; Length ... ;
Best Local Similarity ... ; Pred. No. ... ;
Matches ... Conservative ... Mismatches ... Indels ... ;

Qy ... IIENVHAH ...

Db ... IIENVHAH ...

AAM...
ID AAM... standard; protein; ... AA.
XX
AC AAM...
XX
DT ... (first entry)

XX
KW Human; serine protease; dipeptidyl peptidase IV related protein; DPRP; [illegible]
KW DPPIV; infection; human immunodeficiency virus; HIV 1; HIV 2; [illegible]
KW [illegible]
KW [illegible]
KW [illegible]
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW [illegible]
XX
[illegible]
XX
PT New dipeptidyl peptidase IV related proteins [illegible]
XX
CC The present invention relates to [illegible] human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV) related
CC proteins (DPRP). The dipeptidyl peptidase IV related proteins (DPRP)
CC [illegible]
CC [illegible]
CC [illegible] human immunodeficiency virus (HIV 1 or HIV 2),
CC [illegible] diabetes, [illegible] obesity, anorexia,
CC bulimia, Parkinson's disease, acute heart failure, hypotension,
CC hypertension, urinary retention, osteoporosis, angina pectoris,
CC [illegible]
CC [illegible]
CC [illegible]
CC [illegible] for the treatment of reproductive, inflammatory and
CC [illegible]
XX
SQ Sequence [illegible] AA;

XX
KW Human; [illegible] gastrointestinal; [illegible]
KW cardiovascular; atherosclerosis; autoimmune disorder; [illegible]
KW inflammatory; [illegible] acquired immunodeficiency syndrome; AIDS;
KW [illegible] developmental; [illegible]
KW [illegible] epithelial; [illegible] neurological; [illegible]
KW reproductive; [illegible] endometriosis.
XX
[illegible]
XX
PA [illegible]
XX
PI [illegible]
PI [illegible] AM, [illegible] MB, [illegible] EA, [illegible] A, [illegible]
PI [illegible] NK, [illegible] AM, [illegible] YE, [illegible]
PI [illegible] X, Y, [illegible]
PI [illegible]
XX
DR WPI; [illegible]
DR N-PSDB; [illegible]
XX
PT Twenty [illegible] (referred to as PRTS [illegible]
PT in the diagnosis, treatment and prevention of gastrointestinal (e.g.
PT gastritis), [illegible] cardiovascular (e.g. atherosclerosis) and cell
PT proliferative (e.g. cancer) disorders
XX
PS [illegible]
XX
CC The present invention relates to twenty one new human proteases,
CC referred to as PRTS 1 [illegible] PRTS 21. The PRTS polynucleotides and
CC polypeptides of the invention are useful in the diagnosis, treatment and
CC prevention of gastrointestinal e.g. gastritis, esophageal [illegible]
CC Crohn's disease, cardiovascular e.g. atherosclerosis, hypertension and
CC myocardial infarction, autoimmune/inflammatory e.g. acquired
CC immunodeficiency syndrome (AIDS), allergies and osteoarthritis, [illegible]
CC proliferative [illegible] developmental e.g. Duchenne and Becker
CC muscular dystrophy, epithelial e.g. dermatitis, neurological e.g.
CC epilepsy and Alzheimer's disease and reproductive e.g. infertility [illegible]
CC endometriosis. Numerous other examples of each disorder are
CC [illegible] the present protein sequence represents
CC [illegible] of the invention.
XX
SQ Sequence [illegible]

```
SQ   [illegible] ENVHEAH [illegible]

Db   [illegible] ENVHEAH [illegible]

RESULT [illegible]
AAB4[illegible]
ID   AAB4[illegible] standard; [illegible] AA.
XX
AC   AAB4[illegible];
XX
DT   [illegible] N 2001 (first entry)
XX
DE   Human [illegible].
XX
KW   Human; dipeptidyl aminopeptidase; [illegible] peptidase;
KW   dipeptidyl peptidase; [illegible] IV; [illegible]
KW   growth hormone deficiency; glucose [illegible] mucosa[illegible]
KW   non-insulin dependent diabetes mellitus; [illegible]
KW   immunosuppression.
XX
OS   Homo sapiens.
XX
FH   Key                 Location/Qualifiers
FT   Active-site         [illegible]
FT                       /note= "Forms part of the Asp His catalytic triad"
FT   Active-site         [illegible]
FT                       /note= "Forms part of the Asp His catalytic triad"
FT   Active-site         [illegible]
FT                       /note= "Forms part of the Asp His catalytic triad"
XX
PN   W[illegible] A1.
XX
PD   [illegible] MAR [illegible]
XX
PF   [illegible] W[illegible]AU[illegible].
XX
PR   [illegible] AU[illegible].
PR   [illegible] AU[illegible].
XX
PA   (UNSY ) UNIV SYDNEY.
XX
PI   Abb[illegible] MD;
XX
DR   WPI; [illegible].
DR   N-PSDB; AA[illegible].
XX
PT   New human dipeptidyl aminopeptidase (DPP8) useful [illegible]
PT   substrates, identifying inhibitors of DPP8 catalytic activity which
PT   have therapeutic uses, and for detecting [illegible] cells
XX
PS   Claim [illegible]; Fig [illegible]; 78pp; English.
XX
CC   This sequence represents human dipeptidyl [illegible] peptidase [illegible]
CC   [illegible] has substrate specificity for H-Ala-Pro-pNA, H-Gly-Pro-pNA [illegible]
CC   H-Arg-Pro-pNA. [illegible] peptidase and a
CC   dipeptidyl peptidase, because [illegible]
CC   peptide bond C-terminal to Pro in each of these compounds. [illegible]
CC   [illegible] with human [illegible] DPP8 [illegible]
CC   substrates, and for detecting [illegible]
CC   monitoring the level of DPP8 gene expression [illegible]
CC   [illegible] expression is detected by [illegible]
CC   in the cell. It is also useful for identifying [illegible]
CC   [illegible] the cleavage of the substrate by DPP8. Molecules
CC   identified as inhibitors [illegible] may be used [illegible]
CC   treating diarrhoea, growth hormone deficiency, [illegible]
CC   non-insulin dependent diabetes mellitus and other disorders
CC   [illegible]
CC   as immunosuppressants.
```

```
XX
SQ   [illegible]

  [illegible]

Qy   [illegible]

Db   [illegible]

RESULT [illegible]
AAB4[illegible]
ID   [illegible]
XX
AC   AAB4[illegible]
XX
DT   [illegible]
XX
DE   [illegible]
XX
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
XX
OS   Homo sapiens.
XX
PN   W[illegible]
XX
PD   [illegible]
XX
PF   [illegible]
XX
PR   [illegible]
XX
PA   [illegible]
XX
PI   [illegible]
XX
DR   W[illegible]
DR   N-PSDB; [illegible]
XX
PT   New [illegible]
PT   [illegible]
PT   [illegible]
PT   [illegible]
XX
PS   [illegible]
XX
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
XX
SQ   [illegible]

  [illegible]

Qy   [illegible]

Db   [illegible]

RESULT [illegible]
AAB4[illegible]
ID   [illegible]
XX
```

XX
[illegible]
XX

[illegible]

XX
[illegible]
XX

CC [illegible] novel human [illegible]. These were isolated from expressed sequence [illegible]
CC [illegible]
CC [illegible] to treat aplastic anaemia, to help [illegible]
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple [illegible] to regulate activin or inhibin e.g. to treat
CC infertility [illegible] regulate haemostasis or thrombolysis e.g. to treat
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC [illegible] arthritis, and to treat nervous system disorders [illegible]
CC [illegible] The present sequence is a protein of the [illegible]
XX
SQ [illegible]

[illegible]

XX
KW Human [illegible]
KW [illegible]
KW [illegible]
KW expressed sequence tag.
XX
[illegible]
XX
FA (HYSE-) HYSEQ INC.
XX
[illegible]
XX
CC The present [illegible] provides the protein and [illegible]
CC novel human [illegible]. These were isolated from expressed sequence [illegible]
CC [illegible]
CC [illegible] to treat aplastic anaemia, to help tissue [illegible]
CC [illegible] regulate the immune system [illegible]
CC [illegible] regulate activin or inhibin [illegible]
CC infertility [illegible] regulate haemostasis or thrombolysis [illegible]
CC stroke and [illegible] to screen for drugs, to treat inflammatory [illegible]
CC [illegible] and to treat nervous system disorders [illegible]
CC [illegible] The present sequence is a protein of the [illegible]
XX
SQ [illegible]

```
     [illegible]
     identified as inhibiting [illegible] may be useful for
     treating [illegible] growth hormone deficiency [illegible]
     [illegible] dependent diabetes mellitus and other disorders
     [illegible] glucose intolerance, enhancing [illegible] regeneration and
     as immunosuppressants.
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible] Score [illegible] Length [illegible];
  Best Local Similarity [illegible] Pred. No. [illegible];
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy        [illegible] LLENVHEAH [illegible]

Db       [illegible] LLENVHEAH 407


[illegible]
AAB[illegible]
ID   AAB[illegible] standard; protein; [illegible] AA.
XX
AC   AAB[illegible];
XX
DT   [illegible] MAR [illegible] (first entry)
XX
DE   Human protease PRTS-2 protein.
XX
KW   Human protease PRTS-2; [illegible]
KW   [illegible] dermatitis; arteriosclerosis; [illegible] arthritis; [illegible]
KW   atherosclerosis; psoriasis; Alzheimer's disease; mental disorder; [illegible]
KW   gastrointestinal disorder; Cushing's syndrome; [illegible]
KW   [illegible] disorder; [illegible]
KW   hypertension; neurological disorder; Parkinson's disease; [illegible]
KW   [illegible]; cell proliferative disorder; multiple sclerosis; [illegible]
KW   diabetes mellitus; [illegible] disorder; [illegible]
KW   autoimmune disorder; inflammatory disorder; [illegible]
KW   developmental disorder; reproductive disorder; infertility; [illegible]
KW   [illegible] syndrome; [illegible] epilepsy; [illegible]
XX
OS   Homo sapiens.
XX
PN   WO[illegible]-A2.
XX
PD   [illegible] NOV [illegible]
XX
PF   [illegible] MAY-[illegible] W[illegible].
XX
PR   [illegible] MAY-2000; 2000US-[illegible].
PR   [illegible] MAY-2000; 2000US-[illegible].
PR   [illegible] MAY-2000; 2000US-[illegible].
PR   [illegible] MAY-2000; 2000US-[illegible].
PR   01-JUN-2000; 2000US-[illegible].
XX
PA   (INCY ) INCYTE GENOMICS INC.
XX
PI   [illegible] AM, [illegible], [illegible] A, [illegible], [illegible] NK, [illegible]
PI   [illegible] CM, Khan FA, [illegible] MT, [illegible], [illegible] Y, [illegible]
PI   Nguyen [illegible], [illegible] AR, [illegible], [illegible], [illegible]
PI   Reddy R, [illegible] H, Tang YT;
XX
DR   W[illegible].
DR   N-PSDB; AAC[illegible].
XX
PT   Novel human proteases and polynucleotides encoding the proteases,
PT   useful for treating, diagnosing or preventing cell proliferative,
PT   cardiovascular, autoimmune/inflammatory, neurological and developmental
PT   disorders
XX
PS   [illegible]
XX
CC   The invention relates to human proteases (PRTS-[illegible]) and [illegible]
CC   cDNA [illegible] human [illegible] and [illegible]
```

[illegible]

XX
PA (INT...) [illegible]
XX
XX
New protease (inhibitors) useful for diagnosis and treatment of autoimmune/inflammatory disorders such as acquired immunodeficiency syndrome, Cushing's disease, Addison's disease and cell proliferative disorders such as cancer
XX
[illegible] English
XX
AAB[illegible] to AAB[illegible] encode the human proteases and protease inhibitors (PPIMs) given in AAB74[illegible] to AAB74694. The PPIMs can have activity as: anti-human immunodeficiency virus (HIV); antiinflammatory; anti[illegible]; immunostimulant; immunomodulator; antiinflammatory; immunosuppressive; nephrotropic; antigout; thyromimetic; cytostatic; antibacterial; fungicide; protozoacide; antiarteriosclerotic; antiatherosclerotic; virucide; antipsoriatic; and hepatotropic. PPIM polynucleotide and protein sequences can be used in the diagnosis, treatment and prevention of autoimmune/inflammatory disorders such as AIDS, Addison's disease, severe combined immunodeficiency disease (SCID), [illegible] syndrome, Cushing's disease, Addison's disease, [illegible], Crohn's disease, diabetes mellitus, [illegible] syndrome, [illegible] Grave's disease, Hashimoto's thyroiditis, [illegible] syndrome, [illegible] syndrome, viral, bacterial, fungal, parasitic, protozoal and [illegible] infections and cell proliferative disorder such as arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, psoriasis and cancer. PPIM polynucleotide sequences can be used in somatic or germline gene therapy and in diagnosis of diseases. They can also be used in generating hybridisation probes useful in mapping the naturally occurring gene sequences and in molecular biology techniques.
XX
Sequence [illegible] AA;

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible] Conservative [illegible] Mismatches [illegible]

Qy [illegible] LLENVHEAE [illegible]
Db [illegible] LLENVHEAE [illegible]

[illegible]
AAB[illegible]
ID AAB[illegible] standard; protein; [illegible] AA.
XX
AC AAB[illegible];
XX
DT 24 MAY 2002 (first entry)
XX
DE Human [illegible]
XX
KW [illegible]
KW [illegible]
KW [illegible]
KW [illegible]
KW [illegible] disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
PN WO[illegible]-A2.
XX
PD [illegible] MAY [illegible]
XX

[illegible] MAY [illegible]
XX
PA [illegible]
XX
[illegible]
XX
[illegible]
XX
[illegible] disorders
XX
[illegible]
XX
[illegible]
XX
SQ Sequence [illegible] AA;

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible]

Qy [illegible]

Db [illegible] LLENVHEAE [illegible]

[illegible]
AAB[illegible]
ID AAB[illegible]
XX
AC AAB[illegible]
XX
DT [illegible]
XX
DE Human [illegible]
XX
KW [illegible]
KW [illegible]
KW [illegible]
KW [illegible]
KW [illegible]
XX
OS [illegible]
XX
PN WO[illegible]
XX
PD [illegible] MAY [illegible]
XX
[illegible]
XX
[illegible]
XX
[illegible]
XX

[illegible]

OM protein - protein search, using sw model

Run on: February [illegible], 2003, [illegible]; Search time [illegible] Seconds
(without alignments)
[illegible] Million cell updates/sec

Title: SW464-2
[illegible]: [illegible]
Sequence: [illegible]

[illegible]: [illegible]

Total number of hits satisfying chosen parameters: [illegible]

Minimum DB seq length: [illegible]
Maximum DB seq length: [illegible]

[illegible]: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: A_Geneseq_[illegible]:*
1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA[illegible].DAT:*
[illegible]

SUMMARIES

[illegible]

ALIGNMENTS

RESULT 1
AAR74623
ID AAR74623 standard; protein; [illegible] AA.
XX
AC AAR74623;
XX
DT [illegible] (first entry)
XX
DE Human protease and protease inhibitor PPIM-6.
XX
KW Human; protease; protease inhibitor; protease and protease inhibitor;
KW PPIM; identification; diagnosis; anti human immunodeficiency virus; HIV;
KW antidiabetic; immunostimulant; immunomodulator; antiinflammatory;
KW antithyroid; immunosuppressive; nephrotropic; antigout; thyromimetic;
KW cytostatic; antibacterial; fungicide; protozoacide; antiarteriosclerotic;
KW antiasthmatic; [illegible]; antipsoriatic; virucide; hepatotropic; gene therapy;
KW autoimmune [illegible]; inflammatory disorder; AIDS; Addison's syndrome;
KW severe combined immunodeficiency disease; SCID; Chediak Higashi [illegible];
KW [illegible] disease; Addison's disease; autoimmune thyroiditis; [illegible];
KW Crohn's disease; diabetes mellitus; Goodpasture's syndrome; [illegible];
KW [illegible]; Hashimoto's thyroiditis; Sjogren's syndrome; [illegible];
KW Wegener's [illegible]; cell proliferative disorder; arteriosclerosis;
KW atherosclerosis; [illegible]; hepatitis; psoriasis.
XX
OS Homo sapiens.
XX
PN WO[illegible]-A1.
XX
PD [illegible]
XX
PF [illegible]
XX
PR [illegible]

```
;       STATE: New York
;       COUNTRY: USA
;       ZIP: [illegible]
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Diskette, 3.5 inch, [illegible] MB storage
;       COMPUTER: IBM PS/2
;       OPERATING SYSTEM: PC-DOS
;       SOFTWARE: Wordperfect
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/[illegible]
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]/619,286
;       FILING DATE: 18-MARCH 1996
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: 20-APRIL 1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME: [illegible] No. US2 [illegible]
;       REGISTRATION NUMBER: [illegible],946
;       REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (212) 688-9200
;       TELEFAX: (212) [illegible]
;   INFORMATION FOR SEQ ID NO: 2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: [illegible] amino acids
;       TYPE: amino acid
;       TOPOLOGY: linear
US-[illegible]

  Query Match [illegible]; Score [illegible]; DB [illegible]; Length 7[illegible];
  Best Local Similarity [illegible]; Pred. No. [illegible];
  Matches 7; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy        [illegible] WSYGY [illegible]
              [illegible]
Db        [illegible] WSYGYV [illegible]


RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]
;   GENERAL INFORMATION:
;     APPLICANT: Zimmermann, Rainer; Park, [illegible]
;     APPLICANT: [illegible], Wolfgang; [illegible]
;     TITLE OF INVENTION: [illegible] PLASMINOGEN ACTIVATION [illegible]
;     TITLE OF INVENTION: ALPHA, AND USES THEREOF
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Felfe & Lynch
;       STREET: [illegible] Third Avenue
;       CITY: New York City
;       STATE: New York
;       COUNTRY: USA
;       ZIP: [illegible]
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Diskette, 3.5 inch, [illegible] storage
;       COMPUTER: IBM PS/2
;       OPERATING SYSTEM: PC-DOS
;       SOFTWARE: Wordperfect
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/[illegible]
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: US/[illegible]
;       FILING DATE: 18 MARCH 1996
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: 20-APRIL 1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Hanson, No. US2 [illegible]
```

```
[illegible]
;   INFORMATION FOR SEQ ID NO: [illegible]
[illegible]

Qy        [illegible] WSYGY [illegible]
Db        [illegible] WSYGYV [illegible]

RESULT [illegible]
US [illegible]
; Sequence [illegible], Application [illegible]
; Patent No. [illegible]
;   GENERAL INFORMATION:
;     APPLICANT: [illegible]
;     APPLICANT: [illegible]
;     TITLE OF INVENTION: [illegible]
;     TITLE OF INVENTION: [illegible]
;     NUMBER OF SEQUENCES: [illegible]
;     CURRENT APPLICATION NUMBER: [illegible]
;     PRIOR APPLICATION DATA: [illegible]
;     NUMBER OF SEQ ID NOS: [illegible]
;     SOFTWARE: [illegible]
; SEQ ID NO [illegible]
;   LENGTH: [illegible]
;   TYPE: [illegible]
;   ORGANISM: [illegible]
US-[illegible]

Qy        [illegible]
Db        [illegible]

RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application [illegible]
; Patent No. [illegible]
;   GENERAL INFORMATION:
;     APPLICANT: [illegible]
;     TITLE OF INVENTION: [illegible]
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;       [illegible]
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: [illegible]
;       COMPUTER: [illegible]
```

```
; FILE REFERENCE: [illegible]
; CURRENT APPLICATION NUMBER: US/[illegible]
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US [illegible]/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: [illegible]
; SOFTWARE: PatentIn version [illegible]
; SEQ ID NO [illegible]
; LENGTH: [illegible]
; TYPE: PRT
; ORGANISM: Homo sapiens
US-[illegible]

  Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
  Best Local Similarity 88.9%; Pred. No. [illegible];
  Matches 8; Conservative 1; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible] HWSYGYL [illegible]
Db [illegible] HWSYGYL [illegible]


RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. US2[illegible]84A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO [illegible]
; FILE REFERENCE: 7066[illegible]
; CURRENT APPLICATION NUMBER: US/09/976,[illegible]
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US [illegible]/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: [illegible]
; SOFTWARE: PatentIn version [illegible]
; SEQ ID NO [illegible]
; LENGTH: [illegible]
; TYPE: PRT
; ORGANISM: Homo sapiens
US-[illegible]

  Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
  Best Local Similarity 88.[illegible]%; Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible] HWSY[illegible]
Db [illegible] HWSY[illegible]


RESULT 4
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. US2[illegible]1584A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO [illegible]
; FILE REFERENCE: 7066[illegible]
; CURRENT APPLICATION NUMBER: US/[illegible]
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: [illegible]
; SOFTWARE: PatentIn version [illegible]
; SEQ ID NO [illegible]
; LENGTH: [illegible]
```

```
; [illegible]
; ORGANISM: [illegible]
US-[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy [illegible]

Db [illegible]


RESULT [illegible]
US-[illegible]
; Sequence [illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
; APPLICANT: [illegible]
; APPLICANT: [illegible]
; APPLICANT: [illegible]
; APPLICANT: [illegible]
; TITLE OF INVENTION: [illegible]
; FILE REFERENCE: [illegible]
; CURRENT APPLICATION NUMBER: [illegible]
; CURRENT FILING DATE: [illegible]
; PRIOR APPLICATION NUMBER: [illegible]
; PRIOR FILING DATE: [illegible]
; NUMBER OF SEQ ID NOS: [illegible]
; SOFTWARE: [illegible]
; SEQ ID NO [illegible]
; LENGTH: [illegible]
; TYPE: [illegible]
; ORGANISM: Homo sapiens
US-[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy [illegible]

Db [illegible]


RESULT [illegible]
US-[illegible]
; Sequence [illegible]
; Patent No. US2[illegible]84A1
; GENERAL INFORMATION:
; APPLICANT: [illegible]
; APPLICANT: Akinsanya, [illegible]
; APPLICANT: [illegible]
; APPLICANT: [illegible]
; TITLE OF INVENTION: NOVEL [illegible]
; FILE REFERENCE: [illegible]
; CURRENT APPLICATION NUMBER: [illegible]
; CURRENT FILING DATE: [illegible]
; PRIOR APPLICATION NUMBER: [illegible]
; PRIOR FILING DATE: [illegible]
; NUMBER OF SEQ ID NOS: [illegible]
; SOFTWARE: PatentIn [illegible]
; SEQ ID NO [illegible]
; LENGTH: [illegible]
; TYPE: [illegible]
; ORGANISM: [illegible]
US-[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy [illegible]
```

```
[illegible]
[illegible] Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
X    NCBI_TaxID [illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN [illegible]; TISSUE=MYCELIUM;
RX   MEDLINE [illegible]; PubMed [illegible];
RA   Beauvais A., Monod M., Wyniger [illegible]
RA   Brakch N., Sarfati J., Bouchara [illegible]
RT   "Dipeptidyl-peptidase IV secreted by Aspergillus fumigatus, [illegible]
RT   pathogenic to humans.";
RL   Infect. Immun. [illegible]
DR   EMBL; U87950; AAC34310.1; -.
DR   [illegible]
DR   InterPro; IPR002469; DPPIV_N_term.
DR   InterPro; [illegible]; Peptidase_S9.
DR   InterPro; IPR00[illegible]; Ser_estrs_site.
DR   Pfam; [illegible]; DPPIV_N_term; 1.
DR   Pfam; [illegible]; Peptidase_S9; 1.
KW   Hydrolase.
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible]

Qy  [illegible] WSYIDY [illegible]

Db  [illegible] WSYIDY [illegible]


RESULT [illegible]
[illegible]
ID   [illegible]   PRELIMINARY;   PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   PUTATIVE DIPEPTIDYL AMINOPEPTIDASE [illegible] (EC 3.4.14.[illegible]).
GN   [illegible]
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID [illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN 972;
RA   [illegible]
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. [illegible]
CC       VACUOLES (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS [illegible] PEPTIDASE FAMILY S9B; ALSO KNOWN AS [illegible]
CC       [illegible] DIPEPTIDASE FAMILY.
DR   EMBL; [illegible]
DR   InterPro; IPR002469; DPPIV_N_term.
DR   InterPro; [illegible]; GASA.
DR   InterPro; [illegible]; MIP_family.
DR   InterPro; [illegible]; Peptidase_S9.
DR   InterPro; [illegible]
DR   InterPro; IPR00[illegible]; Ser_estrs_site.
DR   Pfam; [illegible]; DPPIV_N_term; [illegible].
DR   Pfam; PF02[illegible]; GASA; 1.
DR   Pfam; [illegible]; Peptidase_S9; [illegible].
DR   PROSITE; PS00221; MIP; UNKNOWN_1.
DR   PROSITE; [illegible]
KW   Hypothetical protein; Hydrolase; Dipeptidase; [illegible]
KW   Transmembrane; Glycoprotein; [illegible]
FT   DOMAIN        [illegible]     [illegible]      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      [illegible]     [illegible]      POTENTIAL.
FT   DOMAIN        [illegible]     [illegible]      LUMENAL (POTENTIAL).
FT   ACT_SITE      [illegible]     [illegible]      CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE      [illegible]     [illegible]      CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE      [illegible]
FT   CARBOHYD      [illegible]
FT   CARBOHYD      [illegible]
FT   CARBOHYD      [illegible]
FT   CARBOHYD      [illegible]
FT   CARBOHYD      [illegible]
FT   CARBOHYD      [illegible]
FT   CARBOHYD      [illegible]
FT   CARBOHYD      [illegible]
FT   CARBOHYD      [illegible]
FT   CARBOHYD      [illegible]
FT   CARBOHYD      [illegible]
FT   CARBOHYD      [illegible]
SQ   SEQUENCE [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy  [illegible]

Db  [illegible]


[illegible]
[illegible]
```

[illegible]

```
SQ   SEQUENCE [illegible] AA; [illegible] MW; A1998A[illegible] CRC64;
[illegible]
AC   Q9[illegible]N[illegible];
DT   [illegible] MAR-2001 ([illegible]) [illegible] (Created)
DT   01-MAR-2001 ([illegible]) [illegible] Last sequence update)
DT   [illegible] (Last annotation update)
DE   Dipeptidyl peptidase IV like protein.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; [illegible]
OC   eurosids II; [illegible]ales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=[illegible];
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible] Nakamura [illegible] Sato S., Kaneko T., Asamizu E., [illegible] N.,
RA   Tabata S.;
RL   [illegible] Arabidopsis thaliana chromosome [illegible]
DR   EMBL; AB[illegible]
DR   InterPro; [illegible]
DR   InterPro; [illegible] Peptidase_S9.
DR   InterPro; [illegible]
DR   Pfam; [illegible] N-termin[illegible].
DR   Pfam; [illegible]
SQ   SEQUENCE [illegible] MW; [illegible] CRC64;
[illegible]
DE   Dipeptidyl peptidase IV (EC 3.4.14.[illegible]).
OS   Aspergillus [illegible] (Sartorya fumigata).
```

```
OS   [illegible]
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; [illegible]
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Martin L., [illegible];
RT   "The sequence of C. elegans cosmid K[illegible].";
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF[illegible]; AAK[illegible]; [illegible]
DR   InterPro; IPR001375; Peptidase_S9.
DR   InterPro; IPR000379; Ser_estrs_site.
DR   Pfam; PF[illegible]; Peptidase_S9; 1.
SQ   SEQUENCE   927 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match  [illegible]
  Best Local Similarity   88.[illegible];  Pred. No. [illegible];
  Matches  [illegible]  Conservative  [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy   [illegible] HGWSYIDYL [illegible]
Db   [illegible] HGWSYIDYM [illegible]

RESULT [illegible]
[illegible]
ID   [illegible]  PRELIMINARY;  PRT;  [illegible] AA.
AC   [illegible];
DT   [illegible]-[illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   01-MAR-[illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   Hypothetical [illegible] protein.
GN   K[illegible].
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Martin L., [illegible];
RT   "The sequence of C. elegans cosmid K[illegible].";
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF[illegible]; AAK[illegible]; [illegible]
```

```
DR   [illegible]
DR   [illegible]
DR   [illegible]
DR   Pfam; [illegible]
KW   Hypothetical protein.
SQ   SEQUENCE  [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match  [illegible]
  Best Local Similarity  [illegible]
  Matches  [illegible]

Qy   [illegible] HGWSY[illegible]
Db   [illegible] HGWSY[illegible]

RESULT [illegible]
[illegible]
ID   [illegible]
AC   [illegible]
DT   [illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] MAR-[illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   [illegible] cDNA [illegible]
DE   [illegible] IV [illegible]
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; [illegible]
OC   Mammalia; Eutheria; [illegible]
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RA   [illegible]
RA   [illegible]
RA   [illegible] Yamamoto [illegible]
RA   [illegible]
RA   [illegible]
RA   [illegible]
RA   [illegible]
RT   "NEDO [illegible] cDNA [illegible]";
RL   Submitted [illegible] to the EMBL/GenBank/DDBJ databases.
DR   EMBL; [illegible]
DR   InterPro; [illegible]
DR   [illegible]
DR   [illegible]
SQ   SEQUENCE  [illegible]

  Query Match  [illegible]
  Best Local Similarity  [illegible]
  Matches  [illegible]

Qy   [illegible]
Db   [illegible] HGWSY[illegible]

RESULT [illegible]
[illegible]
ID   [illegible]
AC   [illegible]
DT   [illegible]
DT   [illegible]
DT   [illegible]
DE   [illegible]
OS   Homo sapiens (Human).
OC   [illegible]
OC   [illegible]
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   [illegible]
RA   [illegible]
RL   Submitted [illegible] to the EMBL/GenBank/DDBJ databases.
DR   EMBL; [illegible]
```

```
[illegible]
ID   [illegible]     PRELIMINARY;     [illegible] AA.
AC   [illegible]
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   Dipeptidyl peptidase 8 (fragment).
GN   DPP8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE [illegible]; PubMed [illegible];
RA   Abbott C.A., Yu D.M.T., Woollatt E., Sutherland G.R., McCaughan G.W.,
RA   Gorrell M.D.;
RT   "Cloning, expression and chromosomal localization of a novel human
RT   dipeptidyl peptidase (DPP) IV homolog, [illegible].";
RL   Eur. J. Biochem. 267:[illegible] ([illegible]).
DR   EMBL; AF22[illegible]; AAG[illegible].1; -.
DR   MEROPS; S[illegible]; -.
DR   InterPro; IPR00137[illegible]; Peptidase_S9.
DR   InterPro; IPR[illegible]; Ser_estrs_site.
DR   Pfam; PF0[illegible]; Peptidase_S9; 1.
FT   NON_TER [illegible]
FT   NON_TER [illegible]
SQ   SEQUENCE [illegible] AA; 41[illegible] MW; [illegible] CRC64;

  Query Match              100.0%;  Score [illegible];  DB 4;  Length [illegible];
  Best Local Similarity    100.0%;  Pred. No. [illegible];
  Matches    9;  Conservative   [illegible];  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy          1 HGWSYGGY [illegible]
              [illegible]
Db        [illegible] HGWSYGGY [illegible]28


RESULT [illegible]
[illegible]
ID   [illegible]     PRELIMINARY;     PRT;   [illegible] AA.
AC   [illegible]
DT   01-MAR-20[illegible] (TrEMBLrel. 16, Created)
DT   01-MAR-2[illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   01-JUN-2[illegible] (TrEMBLrel. 21, Last annotation update)
DE   Dipeptidyl peptidase 8.
GN   DPP8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE [illegible]; PubMed [illegible];
RA   Abbott C.A., Yu D.M.T., Woollatt E., Sutherland G.R., McCaughan G.W.,
RA   Gorrell M.D.;
RT   "Cloning, expression and chromosomal localization of a novel human
RT   dipeptidyl peptidase (DPP) IV homolog, DPP8.";
RL   Eur. J. Biochem. 267:[illegible].
DR   EMBL; AF22[illegible]; AAG[illegible].1; -.
DR   MEROPS; S[illegible]; -.
DR   InterPro; IPR[illegible]; DPPIV_N_term.
DR   InterPro; IPR001375; Peptidase_S9.
DR   InterPro; IPR[illegible]; Ser_estrs_site.
DR   Pfam; PF[illegible]; DPPIV_N_term; 1.
DR   Pfam; PF[illegible]; Peptidase_S9; 1.
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible] CRC64;

  Query Match              [illegible];  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity    [illegible];  Pred. No. [illegible];
  Matches    [illegible];  Conservative   [illegible];  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy          [illegible] HGWSYGGY [illegible]
```

```
[illegible]
Db        [illegible] HGWSYGGY [illegible]

RESULT [illegible]
[illegible]
ID   [illegible]
AC   [illegible]
DT   [illegible]
DT   [illegible]
DT   [illegible]
DE   [illegible]
GN   [illegible]
OS   [illegible]
OC   Eukaryota; Metazoa; [illegible]
OC   Mammalia; Eutheria; [illegible]
OX   NCBI_TaxID=[illegible]
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   [illegible]
RX   MEDLINE [illegible]; PubMed [illegible]
RA   [illegible]
RT   [illegible]
RL   Nature [illegible]
DR   [illegible]
SQ   SEQUENCE [illegible] AA; [illegible]

  Query Match              [illegible]
  Best Local Similarity    [illegible]
  Matches    [illegible]

Qy          [illegible] HGWSYGGY [illegible]

Db        [illegible] HGWSYGGY [illegible]


RESULT [illegible]
[illegible]
ID   [illegible]     PRELIMINARY;     [illegible]
AC   [illegible]
DT   [illegible] MAY [illegible]
DT   [illegible] MAY [illegible]
DT   [illegible] MAY [illegible]
DE   [illegible]
GN   [illegible]
OS   [illegible]
OC   Eukaryota; Metazoa; Arthropoda; [illegible]
OC   [illegible]
OC   [illegible]
```

```
;        ZIP: 11794
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk, 3.5 inch, 1.44 Mb
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: WordPerfect 5.1 [illegible]
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/ 8/[illegible] 242
;       FILING DATE: 22 December 1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: [illegible] 125/[illegible]
;       FILING DATE: 2[illegible] December 199[illegible]
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Barak, Ronald J.
;       REGISTRATION NUMBER: [illegible]
;       REFERENCE/DOCKET NUMBER: 294 29
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (516) 822 3550
;       TELEFAX: (516) 822 3582
;   INFORMATION FOR SEQ ID NO: 2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 54[illegible] amino acids
;       TYPE: amino acid
;       STRANDEDNESS: unknown
;       TOPOLOGY: linear
;     MOLECULE TYPE: protein
;     HYPOTHETICAL: NO
US-[illegible]

  Query Match [illegible] Score [illegible] Length 54[illegible];
  Best Local Similarity 75.[illegible]%; Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible] WSYTY[illegible]
Db [illegible] WIYGH [illegible]2


RESULT [illegible]
US-08-81[illegible]
; Sequence [illegible], Application US/08814[illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION: Fructosyltransferase Enzymes, Method
;    TITLE OF INVENTION: For Its Production and DNA Encoding the [illegible]
;    NUMBER OF SEQUENCES: 5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Hoffmann & Baron
;      STREET: [illegible] Jericho Turnpike
;      CITY: Jericho
;      STATE: New York
;      COUNTRY: United States of America
;      ZIP: 11753
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk, 3.5 inch, 1.44 Mb
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: WordPerfect 5.1 [illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/ 8/814,1[illegible]
;      FILING DATE: 1[illegible]-MAR-199[illegible]
;      CLASSIFICATION: 435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: US 08/362,242
;      FILING DATE: 22 December 1994
;      APPLICATION NUMBER: GB 125/[illegible]
;      FILING DATE: 2[illegible] December 199[illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Barak, Ronald J.
;      REGISTRATION NUMBER: 2[illegible]
;      REFERENCE/DOCKET NUMBER: 294 29
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: [illegible]
;      TELEFAX: [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: [illegible] amino acids
;      TYPE: [illegible]
;      STRANDEDNESS: unknown
;      TOPOLOGY: [illegible]
;    MOLECULE TYPE: [illegible]
;    HYPOTHETICAL: NO
US-[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy [illegible] WSY[illegible]
Db [illegible]

Search completed: [illegible]
Job time: [illegible]
```

```
[illegible]

; Sequence [illegible], Application [illegible]
; Patent N[illegible]
; GENERAL INFORMATION:
;   APPLICANT: [illegible]
;   APPLICANT: [illegible]
;   TITLE OF INVENTION: [illegible]
;   TITLE OF INVENTION: ALPHA AND USES THEREOF
;   NUMBER OF SEQUENCES: 1
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: [illegible] & Lynch
;     STREET: [illegible] Third Avenue
;     CITY: New York City
;     STATE: New York
;     COUNTRY: USA
;     ZIP: [illegible]
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette, 3.5 inch, [illegible]
;     COMPUTER: IBM PS/2
;     OPERATING SYSTEM: PC-DOS
;     SOFTWARE: Wordperfect
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/[illegible] A
;     FILING DATE: 18 MARCH [illegible]
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/[illegible]
;     FILING DATE: [illegible] APRIL [illegible]4
;   ATTORNEY/AGENT INFORMATION:
;     NAME: [illegible], No. [illegible]
;     REGISTRATION NUMBER: [illegible]
;     REFERENCE/DOCKET NUMBER: [illegible]
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (212) [illegible]
;     TELEFAX: (212) [illegible]
; INFORMATION FOR SEQ ID NO: [illegible]
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: [illegible] amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]; Pred. No. [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]

[illegible] WS[illegible]

[illegible] WS[illegible]


[illegible]
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;   APPLICANT: Zimmermann, [illegible]
;   APPLICANT: Pettit, [illegible]
;   TITLE OF INVENTION: [illegible]
;   TITLE OF INVENTION: PROTEIN ALPHA AND USES THEREOF
;   NUMBER OF SEQUENCES: [illegible]
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: [illegible] & Lynch
;     STREET: [illegible] Third Avenue
;     CITY: New York City
;     STATE: New York
;     COUNTRY: USA
;     ZIP: [illegible]
;   COMPUTER READABLE FORM:
```

[illegible]

```
;      CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: PCT/US93/[illegible]
;        FILING DATE: [illegible]
;        CLASSIFICATION:
;      PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: [illegible]
;        FILING DATE: 21-AUG-1992
;        APPLICATION NUMBER: [illegible]
;        FILING DATE: [illegible]-FEB-1992
;      ATTORNEY/AGENT INFORMATION:
;        NAME: Fraser, Janis K.
;        REGISTRATION NUMBER: [illegible]
;        REFERENCE/DOCKET NUMBER: [illegible]
;      TELECOMMUNICATION INFORMATION:
;        TELEPHONE: (617) 542-5070
;        TELEFAX: (617) 542-8906
;        TELEX: [illegible]154
;   INFORMATION FOR SEQ ID NO: [illegible]:
;      SEQUENCE CHARACTERISTICS:
;        LENGTH: [illegible]
;        TYPE: amino acid
;        STRANDEDNESS:
;        TOPOLOGY: linear
PCT-US93-[illegible]

  Query Match             81.4%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   87.5%;  Pred. No. [illegible];
  Matches [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy        [illegible] WSYFDY [illegible]
              [illegible]
Db        [illegible] WSYFDYV [illegible]24


RESULT [illegible]
PCT-US93-[illegible]
; Sequence 2, Application PC/TUS93[illegible]
;   GENERAL INFORMATION:
;      APPLICANT: M[illegible]
;      APPLICANT: Schlossman, Stuart F.
;      APPLICANT: [illegible]
;      TITLE OF INVENTION: HUMAN [illegible] AND M[illegible]
;      NUMBER OF SEQUENCES: [illegible]
;      CORRESPONDENCE ADDRESS:
;        ADDRESSEE: Fish & Richardson
;        STREET: 225 Franklin Street
;        CITY: Boston
;        STATE: Massachusetts
;        COUNTRY: U.S.A.
;        ZIP: 02110-2804
;      COMPUTER READABLE FORM:
;        MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;        COMPUTER: IBM PS/2 Model 50Z or 55SX
;        OPERATING SYSTEM: IBM P.C. DOS (Version [illegible])
;        SOFTWARE: WordPerfect (Version 5.[illegible])
;      CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: PCT/US93/[illegible]
;        FILING DATE: [illegible]
;        CLASSIFICATION:
;      PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: [illegible]
;        FILING DATE: 21-AUG-1992
;        APPLICATION NUMBER: [illegible]
;        FILING DATE: [illegible]-FEB-1992
;      ATTORNEY/AGENT INFORMATION:
;        NAME: Fraser, Janis K.
;        REGISTRATION NUMBER: [illegible]
;        REFERENCE/DOCKET NUMBER: [illegible]
;      TELECOMMUNICATION INFORMATION:
;        TELEPHONE: (617) 542-5070
;        TELEFAX: (617) 542-[illegible]
;        TELEX: 2[illegible]154
;   INFORMATION FOR SEQ ID NO: [illegible]:
```

```
;        [illegible]
;        LENGTH: [illegible]
;        TYPE: [illegible]
;        STRANDEDNESS: [illegible]
;        TOPOLOGY: [illegible]
[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy        [illegible] WSYFDY [illegible]
              [illegible]
Db        [illegible] WSYFDYV [illegible]


RESULT [illegible]
US-08-[illegible]
; Sequence [illegible], Application [illegible]
; [illegible]
;   GENERAL INFORMATION:
;      APPLICANT: [illegible]
;      APPLICANT: [illegible]
;      TITLE OF INVENTION: [illegible]
;      TITLE OF INVENTION: [illegible]
;      TITLE OF INVENTION: [illegible]
;      NUMBER OF SEQUENCES: [illegible]
;      CORRESPONDENCE ADDRESS:
;        ADDRESSEE: [illegible]
;        STREET: [illegible]
;        CITY: [illegible]
;        STATE: New [illegible]
;        COUNTRY: [illegible]
;        ZIP: [illegible]
;      COMPUTER READABLE FORM:
;        MEDIUM TYPE: [illegible]
;        COMPUTER: [illegible]
;        OPERATING SYSTEM: [illegible]
;        SOFTWARE: [illegible]
;      CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: [illegible]
;        FILING DATE: [illegible]
;        CLASSIFICATION: [illegible]
;      ATTORNEY/AGENT INFORMATION:
;        NAME: [illegible]
;        REGISTRATION NUMBER: [illegible]
;        REFERENCE/DOCKET NUMBER: [illegible]
;      TELECOMMUNICATION INFORMATION:
;        TELEPHONE: [illegible]
;        TELEFAX: [illegible]
;   INFORMATION FOR SEQ ID NO: [illegible]
;      SEQUENCE CHARACTERISTICS:
;        LENGTH: [illegible]
;        TYPE: amino acid
;        [illegible]
[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy        [illegible] WSYFDY [illegible]
              [illegible]
Db        [illegible]


RESULT [illegible]
[illegible]
; Sequence [illegible]
; [illegible]
;   GENERAL INFORMATION:
;      APPLICANT: [illegible]
;      APPLICANT: [illegible]
```

[illegible]

February [illegible] (with [illegible] alignments)

Maximum Match [illegible]

SUMMARIES

[illegible]

ALIGNMENTS

RESULT 1

[illegible]

; Sequence [illegible]
; Patent N[illegible]
; GENERAL INFORMATION
; APPLICANT: [illegible]
; APPLICANT: [illegible]
; APPLICANT: [illegible]
; TITLE OF INVENTION: [illegible]
; FILE REFERENCE: [illegible]
; CURRENT APPLICATION NUMBER: [illegible]
; CURRENT FILING DATE: [illegible]
; NUMBER OF SEQ ID NOS: [illegible]
; SOFTWARE: [illegible]
; SEQ ID NO [illegible]
; LENGTH: [illegible]
; TYPE: [illegible]
; ORGANISM: [illegible]

[illegible]

Best Local Similarity [illegible] Pred. No. [illegible]
Matches [illegible] Mismatches [illegible]

RESULT 2

[illegible]

; Sequence [illegible]
; GENERAL INFORMATION
; APPLICANT: [illegible]
; APPLICANT: [illegible]
; APPLICANT: [illegible]
; TITLE OF INVENTION: [illegible] AND METHOD [illegible]
; NUMBER OF SEQUENCES: [illegible]
; CORRESPONDENCE ADDRESS: [illegible]
; ADDRESSEE: [illegible]
; STREET: [illegible]
; CITY: [illegible]
; STATE: [illegible]
; COUNTRY: [illegible]
; ZIP: [illegible]
; COMPUTER READABLE FORM:
; MEDIUM TYPE: [illegible]
; COMPUTER: [illegible]
; OPERATING SYSTEM: [illegible]
; SOFTWARE: [illegible] (Version [illegible])

```
RESULT 1.
[illegible]
ID   [illegible]   STANDARD;   PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] (Rel. 27, Created)
DT   [illegible] (Rel. 27, Last sequence update)
DT   [illegible] (Rel. [illegible], Last annotation update)
DE   Cuticle protein 7[illegible], isoforms A and B ([illegible]).
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; [illegible]
OC   Insecta; Pterygota; Neoptera; [illegible]
OC   Acridomorpha; Acridoidea; Acrididae; [illegible] Locusta.
OX   NCBI_TaxID=[illegible];
RN   [illegible]
RP   SEQUENCE.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Andersen [illegible], Hoejrup P., Andersen S.[illegible]
RT   "Combined plasma-desorption mass spectrometry and Edman degradation
RT   applied to simultaneous sequence determination [illegible]
RT   structural proteins from the cuticle [illegible] Locusta migratoria.";
RL   Eur. J. Biochem. 217:267-[illegible].
CC   -!- FUNCTION: COMPONENT OF THE CUTICLE [illegible] WHICH
CC       CONTAINS MORE THAN [illegible] DIFFERENT STRUCTURAL PROTEINS.
CC   -!- DOMAIN: THE TETRAPEPTIDE (A-A-P-A/V) REPEATS [illegible]
CC       PROTEIN ARE ALSO PRESENT IN MANY PROTEINS CONSTITUTING THE
CC       PROTECTIVE ENVELOPE OF [illegible] SPECIES.
CC   -!- MISCELLANEOUS: THE SEQUENCE SHOWN HERE IS THAT [illegible]
DR   PIR; S[illegible]; S[illegible].
KW   Structural protein; Cuticle; Repeat.
FT   REPEAT [illegible]
FT   REPEAT [illegible]
FT   REPEAT [illegible]
FT   REPEAT [illegible]
FT   REPEAT [illegible]
FT   VARIANT [illegible] (IN ISOFORM LM-[illegible]).
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible]

  Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
  Best Local Similarity [illegible]; Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

RESULT [illegible]
[illegible]_ARTRM
ID   [illegible]_ARTRM   STANDARD;   PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] (Rel. 4[illegible], Created)
DT   [illegible] (Rel. 4[illegible], Last sequence update)
DT   [illegible] (Rel. 4[illegible], Last annotation update)
DE   Malto-oligosyltrehalose trehalohydrolase [illegible]
DE   alpha-[illegible]-(1->4) alpha-D-glucano-trehalose trehalohydrolase)
DE   (Malto-oligosyltrehalose trehalohydrolase).
GN   [illegible]
OS   Arthrobacter ramosus.
OC   Bacteria; Actinobacteria; [illegible]
OC   Actinomycetales; Micrococcineae; Micrococcaceae; Arthrobacter.
OX   NCBI_TaxID=[illegible];
```

```
RN   [illegible]
RP   SEQUENCE [illegible]
RX   MEDLINE=[illegible]
RA   [illegible]
RT   [illegible]
CC   [illegible]
CC       KNOWN AS [illegible]
CC   [illegible]
CC   between the Swiss [illegible]
DR   EMBL; [illegible]
KW   [illegible]
SQ   SEQUENCE [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

RESULT [illegible]
[illegible]
ID   [illegible]   STANDARD; [illegible]
AC   [illegible]
DT   [illegible]
DE   [illegible]
GN   [illegible]
OS   [illegible]
OC   [illegible]
OX   NCBI_TaxID=[illegible]
RN   [illegible]
RP   SEQUENCE [illegible]
RX   MEDLINE=[illegible]
RA   [illegible]
RT   "Systematic sequencing [illegible]
RL   [illegible]
RN   [illegible]
RP   SEQUENCE [illegible]
RX   MEDLINE=[illegible]
RA   [illegible]
```

[illegible]

```
FT                                      [illegible].
FT   DOMAIN       [illegible]  [illegible]  [illegible] FORM.
FT   DOMAIN       [illegible]  [illegible]  CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     [illegible]  [illegible]  SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
FT                                      (POTENTIAL).
FT   DOMAIN       [illegible]  [illegible]  EXTRACELLULAR (POTENTIAL).
FT   ACT_SITE     [illegible]  [illegible]  CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE     [illegible]  709          CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE     [illegible]  741          CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   CARBOHYD     [illegible]  [illegible]  N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     [illegible]  [illegible]  N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     148          148          N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     217          217          N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     227          227          N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     [illegible]  [illegible]  N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     521          521          N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     686          686          N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     [illegible]  [illegible]  A -> [illegible] (IN REF. [illegible]).
FT   CONFLICT     [illegible]  185          [illegible] -> [illegible] (IN REF. 2).
FT   CONFLICT     [illegible]  [illegible]  [illegible] -> N (IN REF. [illegible]).
FT   CONFLICT     [illegible]  [illegible]  [illegible] -> V (IN REF. [illegible]).
FT   CONFLICT     [illegible]  [illegible]  V -> [illegible] (IN REF. [illegible]).
FT   CONFLICT     [illegible]  [illegible]  [illegible] -> [illegible] (IN REF. [illegible]).
FT   CONFLICT     [illegible]  [illegible]  [illegible] -> [illegible] (IN REF. [illegible]).
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible];  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy    [illegible] WSYTDY [illegible]

Db    [illegible] WSYTDYV [illegible]


RESULT [illegible]
SACB_ACE[illegible]
ID   SACB_ACE[illegible]   STANDARD;   PRT;   [illegible] AA.
AC   [illegible];
DT   [illegible] (Rel. 37, Created)
DT   [illegible] (Rel. 37, Last sequence update)
DT   [illegible] (Rel. [illegible], Last annotation update)
DE   Levansucrase precursor (EC 2.4.1.10) (Beta-D-fructofuranosyl
DE   transferase) (Sucrose 6-fructosyl transferase).
GN   LSDA.
OS   Acetobacter diazotrophicus (Gluconacetobacter diazotrophicus).
OC   Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC   Gluconacetobacter.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF [illegible].
RC   STRAIN=SRT4 / CBS [illegible];
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Arrieta J., Hernandez L., [illegible] A., [illegible]
RA   Menendez C., Petit-Glatron M.-F., Chambert R., Selman-Housein G.;
RL   "Molecular characterization of the levansucrase gene from the
RL   endophytic sugarcane bacterium Acetobacter diazotrophicus SRT4.";
RL   Microbiology 142:1077-1085(1996).
CC   -!- FUNCTION: RELEASES [illegible] SACCHARIDES AND LEVAN [illegible]
CC       MOLECULAR-MASS FRUCTOSYL POLYMERS [illegible]
CC       SUCROSE HYDROLASE THAN AS A FRUCTAN POLYMERASE.
CC   -!- CATALYTIC ACTIVITY: Sucrose + (2,6-beta-D-fructosyl)(N)
CC       = glucose + (2,6)-beta-D-fructosyl(N+1).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PTM: THE N-TERMINAL IS BLOCKED.
CC   -!- SIMILARITY: BELONGS TO FAMILY [illegible] GLYCOSYL HYDROLASES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. [illegible]
```

```
CC   [illegible]
CC   [illegible]
CC
DR   EMBL; [illegible]
DR   [illegible]
DR   [illegible]
KW   [illegible]
FT   [illegible]
FT   [illegible]
FT   DOMAIN [illegible]
SQ   SEQUENCE [illegible] AA; [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy    [illegible] WSYTDY [illegible]

Db    [illegible] WSYTDYV [illegible]


RESULT [illegible]
[illegible]
ID   [illegible]
AC   [illegible]
DT   [illegible]
DT   [illegible]
DT   [illegible]
DE   [illegible]
GN   [illegible]
OS   [illegible]
OC   [illegible]
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RX   [illegible]
RA   [illegible]
RT   [illegible]
RT   [illegible]
RL   [illegible]
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=[illegible]
RA   [illegible]
RT   [illegible]
RT   [illegible]
RL   [illegible]
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=[illegible]
RX   MEDLINE=[illegible]
RA   [illegible]
RT   [illegible]
RT   [illegible]
RL   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   This SWISS-PROT entry is copyright. [illegible]
CC   between the Swiss Institute of Bioinformatics [illegible]
CC   the European Bioinformatics Institute. [illegible]
CC   use by [illegible]
CC   modified and this statement is not removed. [illegible]
CC   [illegible] (See http://www.[illegible]
CC   [illegible]
CC
DR   EMBL; [illegible] AA[illegible]
DR   EMBL; [illegible] AA[illegible]
```

```
DR   EMBL; [illegible] AAB[illegible]; [illegible] DNA.
DR   EMBL; [illegible] AAB[illegible]; [illegible] DNA.
DR   EMBL; [illegible] AAB[illegible]; [illegible] DNA.
DR   EMBL; [illegible] AAB[illegible]; [illegible] DNA.
DR   EMBL; [illegible] AAB[illegible]; [illegible] DNA.
DR   EMBL; [illegible] AAB[illegible]; [illegible] DNA.
DR   EMBL; [illegible] AAB[illegible]; [illegible] DNA.
DR   EMBL; [illegible] AAB[illegible]; [illegible] DNA.
DR   EMBL; [illegible] AAB[illegible]; [illegible] DNA.
DR   EMBL; [illegible] AAB[illegible]; [illegible] DNA.
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAB[illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAB[illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]
DR   MIM; [illegible]
DR   InterPro; [illegible]
DR   InterPro; [illegible]
DR   InterPro; [illegible]
DR   InterPro; [illegible]
DR   Pfam; [illegible]
DR   Pfam; [illegible]
DR   PROSITE; PS[illegible]; PRO_ENDOPEP_SER; 1.
KW   Hydrolase; Aminopeptidase; Dipeptidase; Serine protease;
KW   Transmembrane; Signal-anchor; Glycoprotein; [illegible]
FT   CHAIN        1    [illegible]   DIPEPTIDYL PEPTIDASE IV MEMBRANE FORM
FT                                (MDPP).
FT   CHAIN      [illegible]  [illegible]   DIPEPTIDYL PEPTIDASE IV SOLUBLE FORM
FT                                (SDPP).
FT   DOMAIN     [illegible]  [illegible]   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   [illegible]  [illegible]   SIGNAL ANCHOR (TYPE II MEMBRANE PROTEIN)
FT                                (POTENTIAL).
FT   DOMAIN     [illegible]  [illegible]   EXTRACELLULAR (POTENTIAL).
FT   ACT_SITE   [illegible]  [illegible]   CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   [illegible]  [illegible]   CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   [illegible]  [illegible]   CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   CARBOHYD   [illegible]  [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   [illegible]  [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   [illegible]  [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   [illegible]  [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   [illegible]  [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   [illegible]  [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   [illegible]  [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   [illegible]  [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   [illegible]  [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CONFLICT   [illegible]  [illegible]   [illegible] (IN REF. [illegible]).
FT   CONFLICT   [illegible]  [illegible]   [illegible] (IN REF. [illegible]).
FT   CONFLICT   [illegible]  [illegible]   [illegible] (IN REF. [illegible]).
FT   CONFLICT   [illegible]  [illegible]   [illegible] (IN REF. [illegible]).
FT   CONFLICT   [illegible]  [illegible]   [illegible] (IN REF. [illegible]).
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

[illegible]

[illegible] SWISSPROT [illegible]

[illegible] SWISSPROT [illegible]

[illegible]
[illegible]
[illegible] STANDARD; [illegible] AA.
AC   [illegible]
DT   [illegible] (Rel. [illegible], Created)
DT   [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] (Rel. [illegible], Last annotation update)
DE   Dipeptidyl peptidase IV [illegible]
DE   [illegible]

RP   [illegible]
RN   [illegible]
RC   [illegible]
RX   [illegible]
RA   [illegible]
RL   [illegible]
RN   [1]
RP   SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible] Ikehara Y.;
RT   "[illegible] dipeptidyl peptidase IV [illegible]
RT   the cDNA [illegible] of the NH2-terminal signal sequence [illegible]
RT   the membrane anchoring domain.";
RL   [illegible]
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible]
RT   "[illegible] domain specific membrane
RT   [illegible]";
RL   [illegible] U.S.A. [illegible]
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible]
RT   "[illegible] the NH2-terminal signal/membrane [illegible]
RT   rat dipeptidyl peptidase IV.";
RL   [illegible]
RN   [4]
RP   SEQUENCE [illegible]
RP   [illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible]
RT   "N-terminal [illegible] acid sequence of the [illegible] protein [illegible]
RT   dipeptidyl peptidase IV.";
RL   [illegible]
CC   -!- FUNCTION: REMOVES N-TERMINAL DIPEPTIDES SEQUENTIALLY FROM
CC       POLYPEPTIDES HAVING UNSUBSTITUTED N-TERMINI PROVIDED THAT THE
CC       PENULTIMATE RESIDUE IS PROLINE.
CC   -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC       Xaa, from a polypeptide, preferentially when Xbb is Pro, provided
CC       Xaa is neither Pro nor hydroxyproline.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS IN
CC       A SOLUBLE FORM.
CC   -!- [illegible] FORM (SDPP) DERIVES FROM THE MEMBRANE FORM [illegible]
CC       [illegible]
CC   -!- SIMILARITY: [illegible] FAMILY S9B.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]
DR   EMBL; [illegible]
DR   PIR; [illegible]
DR   MEROPS; [illegible]
DR   InterPro; [illegible] DPPIV_N_term.
DR   InterPro; [illegible] Peptidase_S9.
DR   InterPro; [illegible]
DR   InterPro; [illegible]
DR   Pfam; [illegible]
DR   Pfam; [illegible] DPPIV_N_term; 1.
DR   PROSITE; [illegible] PRO_ENDOPEP_SER; 1.
KW   Hydrolase; Aminopeptidase; Dipeptidase; Serine protease;
KW   Transmembrane; Signal-anchor; [illegible]
FT   CHAIN      [illegible]  [illegible]   DIPEPTIDYL PEPTIDASE IV MEMBRANE FORM
FT                                (MDPP).
FT   CHAIN      [illegible]  [illegible]   DIPEPTIDYL PEPTIDASE IV SOLUBLE FORM
```

```
FT   CARBOHYD    [illegible]    [illegible]       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    [illegible]    [illegible]       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    [illegible]    [illegible]       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    28[illegible]  28[illegible]     N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    [illegible]    [illegible]       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    [illegible]    [illegible]       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    [illegible]    [illegible]       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    [illegible]    [illegible]       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    [illegible]    [illegible]       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   76[illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          [illegible] WSYTHY [illegible]
                [illegible]
Db      [illegible] WSYTHY [illegible]


RESULT [illegible]
DPP4_HUMAN
ID   DPP4_HUMAN     STANDARD;      PRT;   76[illegible] AA.
AC   [illegible];
DT   [illegible]-1992 (Rel. 2[illegible], Created)
DT   [illegible]-FEB-1996 (Rel. [illegible], Last sequence update)
DT   [illegible]-JUN-2002 (Rel. 41, Last annotation update)
DE   Dipeptidyl peptidase IV (EC 3.4.14.[illegible]) (DPP IV) (T-cell activation
DE   antigen CD26) (TP103) (Adenosine deaminase complexing protein 2)
DE   (ADABP).
GN   DPP4 OR ADCP2 OR CD26.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=9232[illegible]; PubMed=[illegible];
RA   Misumi Y., Hayashi Y., Arakawa F., Ikehara Y.;
RT   "Molecular cloning and sequence analysis of human dipeptidyl
RT   peptidase IV, a serine proteinase on the cell surface.";
RL   Biochim. Biophys. Acta [illegible] (199[illegible]).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=95012454; PubMed=[illegible];
RA   Abbott C.A., Baker E., Sutherland G.R., McCaughan G.W.;
RT   "Genomic organization, exact localization, and tissue expression of
RT   the human CD26 (dipeptidyl peptidase IV) gene.";
RL   Immunogenetics 40:331-338(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Peripheral blood;
RX   MEDLINE=92325476; PubMed=[illegible];
RA   Tanaka T., Camerini D., Seed B., Torimoto Y., Dang N.H., Kameoka J.,
RA   Dahlberg H.N., Schlossman S.F., Morimoto C.;
RT   "Cloning and functional expression of the T cell activation antigen
RT   CD26.";
RL   J. Immunol. 149:481-486(1992).
RN   [4]
RP   [illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Tanaka T.;
RL   J. Immunol. [illegible](199[illegible]).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Prostate;
RA   Strausberg R.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE [illegible] FROM N.A.
RC   TISSUE=[illegible];
```

```
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible]
RA   [illegible]
RT   "[illegible]
RT   [illegible]
RT   [illegible]
RT   [illegible]"
RL   [illegible]
RN   [illegible]
RP   SEQUENCE [illegible]
RC   [illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible]
RT   "[illegible]
RT   dipeptidyl peptidase [illegible]
RT   [illegible]"
RL   [illegible]
RN   [illegible]
RP   [illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible]
RT   "[illegible]
RT   [illegible]
RT   [illegible]"
RL   [illegible]
RN   [illegible]
RP   [illegible]
RC   [illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible]
RA   [illegible]
RT   "[illegible]
RT   [illegible]"
RL   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   This SWISS-PROT entry [illegible]
CC   between the Swiss [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
DR   EMBL; [illegible]; AAA[illegible]
DR   EMBL; [illegible]; AAA[illegible]
DR   EMBL; [illegible]; AAA[illegible]
DR   EMBL; [illegible]; AAA[illegible]
DR   EMBL; [illegible]; AAA[illegible]
DR   EMBL; [illegible]; AAA[illegible]
DR   EMBL; [illegible]; AAA[illegible]
DR   EMBL; [illegible]; AAA[illegible]
DR   EMBL; [illegible]; AAA[illegible]
DR   EMBL; [illegible]; AAA[illegible]
DR   EMBL; [illegible]; AAA[illegible]
DR   EMBL; [illegible]; AAA[illegible]
DR   EMBL; [illegible]; AAA[illegible]
DR   EMBL; [illegible]; AAA[illegible]
DR   EMBL; [illegible]; AAA[illegible]
```

```
RX   MEDLINE [illegible] PubMed [illegible]
RT   "[illegible] adenosine deaminase [illegible] protein from [illegible]
RT   kidney [illegible] identical to human [illegible] and rat dipeptidyl peptidase
RT   IV."
RL   [illegible] Physiol. [illegible]
CC   -!- FUNCTION: Removes N-terminal dipeptides sequentially from
CC       polypeptides having unsubstituted N-termini provided that the
CC       penultimate residue is proline. Binds and regulates the activity
CC       of ADA.
CC   -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Xbb-
CC       [illegible] from a polypeptide, preferentially when Xbb is Pro, provided
CC       [illegible]
CC   -!- [illegible] homodimer.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein. [illegible]
CC       [illegible] (By similarity).
CC   -!- [illegible]
CC       [illegible]
CC   -!- [illegible] form (SDPP) derives from the membrane form (MDPP)
CC       [illegible] (By similarity).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S9B.
CC
CC   [illegible]
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. [illegible]
CC   [illegible]
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://[illegible]
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; AF[illegible]; AAF[illegible]
DR   EMBL; AY[illegible]; AA[illegible]
DR   MEROPS; S09.003; -.
DR   InterPro; IPR[illegible]; DPPIV_N_term.
DR   InterPro; IPR[illegible]; Peptidase_S9.
DR   InterPro; IPR[illegible]; Prol_endopep_S9.
DR   InterPro; IPR[illegible]
DR   Pfam; PF[illegible]; DPPIV_N_term; 1.
DR   Pfam; PF00326; Peptidase_S9; 1.
DR   PROSITE; [illegible]
KW   Hydrolase; Aminopeptidase; Dipeptidyl peptidase; Serine protease;
KW   Transmembrane; Glycoprotein; Signal-anchor.
FT   CHAIN        [illegible]     DIPEPTIDYL PEPTIDASE IV MEMBRANE FORM
FT                                (MDPP).
FT   CHAIN        [illegible]     DIPEPTIDYL PEPTIDASE IV SOLUBLE FORM
FT                                (SDPP) (BY SIMILARITY).
FT   DOMAIN       [illegible]     CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     [illegible]     SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
FT                                (POTENTIAL).
FT   DOMAIN       [illegible]     EXTRACELLULAR (POTENTIAL).
FT   ACT_SITE     [illegible]     CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE     [illegible]     CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE     [illegible]     CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   CARBOHYD     [illegible]     N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     [illegible]     N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     [illegible]     N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     [illegible]     N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     [illegible]     N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     [illegible]     N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     [illegible]     N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     [illegible]     N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     [illegible]     N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     [illegible]     N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

[illegible]

RESULT [illegible]
DPP4_FELCA
ID   DPP4_FELCA     STANDARD;      PRT;   [illegible] AA.
AC   Q9N2I7;
DT   [illegible] (Rel. [illegible], Created)
DT   [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] (Rel. [illegible], Last annotation update)
DE   Dipeptidyl peptidase IV (EC 3.4.14.5) (DPP IV) [illegible]
DE   antigen CD26).
GN   DPP4 [illegible]
OS   [illegible]
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RC   [illegible]
RX   MEDLINE [illegible] PubMed [illegible]
RA   Nishimori Y., [illegible] Ikeda Y., [illegible] Y., Nakamura K., [illegible]
RA   Mikami [illegible]
RT   "Molecular [illegible] and sequencing of a cDNA encoding the feline [illegible]
RT   [illegible] CD26 from [illegible]."
RL   [illegible]
CC   -!- FUNCTION: Removes N-terminal dipeptides sequentially from
CC       polypeptides having unsubstituted N-termini provided that the
CC       penultimate residue is proline (By similarity).
CC   -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Xbb-
CC       [illegible] from a polypeptide, preferentially when Xbb is Pro, provided
CC       [illegible] hydroxyproline.
CC   -!- SUBUNIT: Homodimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein. Also exists [illegible]
CC       [illegible] (By similarity).
CC   -!- PTM: The soluble form (SDPP) derives from the membrane form (MDPP)
CC       by proteolytic processing (By similarity).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S9B.
CC
CC   [illegible]
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions [illegible]
CC   [illegible]
CC   modified and this statement is not removed. [illegible]
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   [illegible]
CC
DR   EMBL; AB[illegible]
DR   MEROPS; [illegible]
DR   InterPro; [illegible] DPPIV_N_term.
DR   InterPro; IPR[illegible]; Peptidase_S9.
DR   InterPro; [illegible]
DR   InterPro; [illegible]
DR   Pfam; PF[illegible]; DPPIV_N_term; 1.
DR   Pfam; [illegible] Peptidase_S9; 1.
DR   PROSITE; [illegible]
KW   Hydrolase; Aminopeptidase; Dipeptidyl peptidase; Serine protease;
KW   Transmembrane; Glycoprotein; Signal-anchor.
FT   CHAIN        [illegible]     DIPEPTIDYL PEPTIDASE IV MEMBRANE FORM
FT                                (MDPP).
FT   CHAIN        [illegible]     DIPEPTIDYL PEPTIDASE IV SOLUBLE FORM
FT                                (SDPP) (BY SIMILARITY).
FT   DOMAIN       [illegible]     CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     [illegible]     SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
FT                                (POTENTIAL).
FT   DOMAIN       [illegible]     EXTRACELLULAR (POTENTIAL).
FT   ACT_SITE     [illegible]     CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE     [illegible]     CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE     [illegible]     CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   CARBOHYD     [illegible]     N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     [illegible]     N-LINKED (GLCNAC...) (POTENTIAL).
```

```
RESULT 4
SEPR_MOUSE
ID   SEPR_MOUSE     STANDARD;      PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible]-JUN-2002 (Rel. 41, Created)
DT   [illegible]-JUN-2002 (Rel. 41, Last sequence update)
DT   [illegible]-JUN-2002 (Rel. 41, Last annotation update)
DE   Seprase (EC [illegible]) (Fibroblast activation protein alpha) (Integral
DE   membrane serine protease).
GN   FAP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS [illegible] AND [illegible]).
RC   STRAIN=BALB/c; TISSUE=Embryo;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Niedermeyer J., Scanlan M.J., Garin-Chesa P., [illegible]
RA   [illegible] L.J., Rettig W.J., Schnapp A.;
RT   "Mouse fibroblast activation protein: molecular cloning, alternative
RT   splicing and expression in the reactive stroma of epithelial
RT   cancers.";
RL   Int. J. Cancer [illegible]
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM [illegible]).
RC   TISSUE=Breast;
RA   Strausberg R.;
RL   Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: May have a role in tissue remodeling during development
CC       and wound healing, and contribute to invasiveness in malignant
CC       cancers.
CC   -!- CATALYTIC ACTIVITY: Degrades gelatin and heat denatured type I and
CC       type IV collagen, but not native type I or type IV collagen, [illegible]
CC       [illegible] laminin, fibronectin, fibrin or casein.
CC   -!- SUBUNIT: Homodimer, [illegible] with [illegible]
CC       inactive (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein (By similarity).
CC   -!- ALTERNATIVE PRODUCTS: [illegible] (shown below), [illegible] and [illegible]
CC       produced by alternative splicing.
CC   -!- TISSUE SPECIFICITY: Detected in fibroblasts, in placenta, [illegible]
CC       embryos from day 7 [illegible] and in new born [illegible]
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S9B.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. [illegible]
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Y[illegible]; CAA[illegible].1; -.
DR   EMBL; BC0[illegible]; AAH[illegible].1; -.
DR   MGD; MGI:[illegible]; Fap.
DR   InterPro; IPR[illegible]; DPPIV_N_term.
DR   InterPro; IPR[illegible]; Peptidase_S9.
DR   InterPro; IPR[illegible]; [illegible]
DR   InterPro; IPR[illegible]; Ser_[illegible]
DR   Pfam; PF[illegible]; Peptidase_S[illegible]
```

```
DR   [illegible]
DR   Pfam; PF00326; Peptidase_S9; 1.
DR   [illegible] DPPIV_N_term; 1.
DR   [illegible] PRO_ENDOPEP_SER; 1.
KW   Hydrolase; Aminopeptidase; Dipeptidase; Serine protease;
KW   Transmembrane; Glycoprotein; Signal-anchor.
FT   [illegible]MAIN    1     [illegible]   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   [illegible]   4[illegible]   SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN).
FT   [illegible]MAIN    [illegible]   818   LUMENAL (POTENTIAL).
FT   ACT_SITE   [illegible]   [illegible]   CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   [illegible]   75[illegible]   CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   [illegible]   78[illegible]   CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   CARBOHYD   [illegible]   [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   [illegible]   [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   [illegible]   [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   [illegible]   [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   [illegible]72   [illegible]72   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   [illegible]   [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   [illegible]   42[illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   [illegible]   [illegible]   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CONFLICT   83   83   Q -> H (IN REF. 1).
FT   CONFLICT   [illegible]   125   S -> N (IN REF. 1).
FT   CONFLICT   [illegible]   1[illegible]   [illegible] (IN REF. [illegible]).
FT   CONFLICT   [illegible]   2[illegible]   [illegible] -> N (IN REF. [illegible]).
FT   CONFLICT   [illegible]   375   ISNVVRNISS -> [illegible]FKRIKRRKF (IN REF. 1).
FT   CONFLICT   [illegible]   818   AKRAFL[illegible]PVK -> [illegible]SVISMINLINF[illegible]IYSSYHRL[illegible]
FT                                 FSYLHRRYI (IN REF. 1).
SQ   SEQUENCE   8[illegible] AA;  9[illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches    8;  Conservative    0;  Mismatches    [illegible];  Indels    [illegible];  Gaps    [illegible]

Qy          [illegible] WSY[illegible]
Db        677 WSYGGYL 684


RESULT 2
DPP4_MOUSE
ID   DPP4_MOUSE     STANDARD;      PRT;   [illegible] AA.
AC   [illegible]
DT   01-DEC-1992 (Rel. 24, Created)
DT   [illegible] FEB 199[illegible] (Rel. 3[illegible], Last sequence update)
DT   [illegible] 2[illegible] (Rel. 4[illegible], Last annotation update)
DE   Dipeptidyl peptidase IV (EC 3.4.14.[illegible]) [illegible]
DE   antigen CD26) (Thymocyte activating molecule) (THAM).
GN   DPP4 [illegible]
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Swiss; TISSUE=Thymus;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Marguet D.A., Bernard A.M., Vivier [illegible] Naquet [illegible]
RA   Pierres M.;
RT   "cDNA cloning for mouse thymocyte activating molecule. A
RT   [illegible] dipeptidyl peptidase IV (CD26) [illegible]
RT   [illegible] serine proteases.";
RL   J. Biol. Chem. 267:220[illegible]-2208(1992).
RN   [2]
RP   REVISIONS.
RA   Marguet D.A.;
RL   Submitted (SEP-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=[illegible]; TISSUE=Liver;
RX   MEDLINE=[illegible]2780; PubMed=[illegible];
RA   Bernard A.M., Mattei M.[illegible], Pierres M., Marguet [illegible]
RT   "Structure of the mouse dipeptidyl peptidase IV (CD26) gene."
RL   Biochemistry [illegible]:[illegible]204-[illegible](1994).

RN   [illegible]
RP   [illegible]
RA   [illegible]
RL   [illegible]
RN   [illegible]
RP   SEQUENCE [illegible]
RX   MEDLINE[illegible]
RA   [illegible]
RA   Bernard A.M. [illegible]
RT   [illegible]
CC   -!- FUNCTION: [illegible]
CC   -!- CATALYTIC ACTIVITY: [illegible]
CC   -!- SUBCELLULAR LOCATION: [illegible]
CC   -!- SIMILARITY: [illegible]
CC   [illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   EMBL; [illegible] AAA[illegible]
DR   [illegible]
KW   [illegible]
KW   [illegible]
FT   CHAIN      [illegible]
FT   [illegible]
FT   TRANSMEM   [illegible]
FT   [illegible]
```

SUMMARY

No.	Score	Match	Length	DB	ID	Description
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	SELE_HUMAN	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	VAVL_HUMAN	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	WNCA_HUMAN	[illegible]

[illegible]

ALIGNMENTS

RN [2]
RX MEDLINE [illegible] PubMed [illegible]
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., [illegible]
RA [illegible] Fulton [illegible] Gattung S., Geisel [illegible] Kirsten [illegible]
RA [illegible] Johnston [illegible] Langston Y.,
RA [illegible] Martin [illegible] Meneses S., [illegible]
RA Nhan M., [illegible] St Peter H., Trevaskis E., Vaughan K.
RA Vignati [illegible] Wohldman P., Waterston R., Wilson R.,
RA Vaudin M.;
RL Science [illegible] (1994).
CC [illegible] TYPE II MEMBRANE PROTEIN, LYSOSOME [illegible]
CC [illegible] FAMILY [illegible]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. [illegible]
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
DR EMBL; [illegible]
DR EMBL; [illegible]

[illegible]

[illegible]
[illegible]
[illegible] peptidase family protein (imported) [illegible]
C;Species: [illegible]
C;Date: [illegible] Apr [illegible] #sequence_revision [illegible] Apr [illegible] #text_change [illegible] Apr [illegible]
C;Accession: [illegible]
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, [illegible]
R.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; [illegible]
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, [illegible]
Proc. Natl. Acad. Sci. U.S.A. [illegible] 4136-4141, [illegible]
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:[illegible]; PMID:[illegible]
A;Accession: [illegible]
A;Status: preliminary
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: [illegible] N[illegible] [illegible] N:AAK[illegible]
C;Genetics:
A;Gene: [illegible]

Query Match [illegible]; Score [illegible]; [illegible] Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible] WSY[illegible]
[illegible]
Db [illegible] WSY[illegible] 486

[illegible]
[illegible]
probable dipeptidyl aminopeptidase - fission yeast (Schizosaccharomyces [illegible]
C;Species: Schizosaccharomyces pombe
C;Date: [illegible]-Sep-[illegible] #sequence_revision [illegible]-Sep-[illegible] #text_change [illegible]
C;Accession: [illegible]
R;Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August [illegible]
A;Reference number: Z21737
A;Accession: [illegible]
A;Status: preliminary; translated from GB/EMBL/[illegible]
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: EMBL:[illegible]; PIDN:[illegible]; GSPDB:GN[illegible]
A;Experimental source: strain [illegible]; cosmid [illegible]
C;Genetics:
A;Gene: SPB[illegible]
A;Map position: [illegible]

Query Match [illegible]; Score [illegible]; [illegible] Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible] WSY[illegible]
[illegible]
Db [illegible] WSY[illegible] 614

[illegible]
[illegible]
hypothetical protein C27C12.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: [illegible] #sequence_revision [illegible] #text_change [illegible]
C;Accession: [illegible]
R;Thomas, K.
submitted to the EMBL Data Library, March [illegible]
A;Reference number: Z19134
A;Accession: [illegible]
A;Status: preliminary; translated from GB/EMBL/[illegible]
A;Molecule type: DNA
A;Residues: [illegible] W[illegible]
A;Cross-references: EMBL:[illegible]; PIDN:AA[illegible]; GSPDB:GN[illegible]
A;Experimental source: clone C27C12

[illegible]
A;Map [illegible]
A;[illegible]
C;Superfamily: [illegible]

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible]

Qy [illegible] W[illegible]
[illegible]
Db [illegible] W[illegible]

[illegible]
[illegible]
dipeptidyl [illegible]
C;Species: [illegible]
C;Date: [illegible]
C;Accession: [illegible]
R;Nierman, W.C.; [illegible]
R.; Laub, M.T.; [illegible]
n, J.; Ermolaeva, M.; White, [illegible]
Proc. Natl. Acad. Sci. [illegible]
A;Title: [illegible]
A;Reference number: [illegible]
A;Accession: [illegible]
A;Status: preliminary
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: [illegible]
C;Genetics:
A;Gene: [illegible]

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible]

Qy [illegible] WSY[illegible]
[illegible]
Db [illegible] WSY[illegible]

[illegible]
[illegible]
hypothetical protein [illegible]
C;Species: [illegible]
C;Date: [illegible]
C;Accession: [illegible]
R;W[illegible]
submitted to the EMBL Data Library, [illegible]
A;Reference number: [illegible]
A;Accession: [illegible]
A;Status: [illegible]
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: EMBL:[illegible]
A;Experimental source: [illegible]
C;Genetics:
A;Gene: [illegible]
A;Map position: [illegible]
A;Introns: [illegible]
C;Superfamily: [illegible]

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible]

Qy [illegible] WSY[illegible]
[illegible]
Db [illegible] WSY[illegible]

```
Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy [illegible] WSYYDN [illegible]

Db [illegible] WSYYDY [illegible]

[illegible]
A[illegible]
dipeptidyl aminopeptidase B [illegible] yeast (Saccharomyces [illegible])
N;Alternate names: protein YHR[illegible]
C;Species: Saccharomyces cerevisiae
C;Date: [illegible] Jun 1996 #sequence_revision [illegible] May [illegible] #text_change [illegible]
C;Accession: S[illegible]; A[illegible]
[illegible]
submitted to the EMBL Data Library, June 1994
A;Description: The sequence of [illegible] cerevisiae [illegible]
A;Reference number: S[illegible]
A;Accession: S[illegible]
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: EMBL:[illegible]; NID:[illegible]; PIDN:AA[illegible]; PID:[illegible]; SGD:[illegible]
R;Roberts, C.J.; Pohlig, G.; Rothman, J.H.; Stevens, T.H.
J. Cell Biol. [illegible], 1989
A;Title: Structure, biosynthesis, and localization of dipeptidyl aminopeptidase B, an [illegible]
A;Reference number: A[illegible]; MUID:[illegible]; PMID:[illegible]
A;Accession: A[illegible]
A;Molecule type: DNA
A;Residues: [illegible],'H',84-124,'N',1[illegible]-181 [illegible],'N',[illegible]
A;Cross-references: EMBL:X1[illegible]
A;Note: the authors translated the codon A[illegible] for residue [illegible] as As[illegible]
C;Genetics:
A;Gene: SGD:DAP2; STE13; MIPS:YHR[illegible]
A;Cross-references: SGD:[illegible]; MIPS:YHR[illegible]
A;Map position: [illegible]
C;Superfamily: dipeptidyl-peptidase IV
C;Keywords: dipeptidylpeptidase hydrolase; [illegible] transmembrane [illegible]
F;[illegible]/Domain: transmembrane #status predicted [illegible]
F;[illegible]/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
  Best Local Similarity [illegible]; Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible] WSYYDN [illegible]

Db [illegible] WSYYDY [illegible]

[illegible]
S[illegible]
X-[illegible] dipeptidyl-peptidase (EC 3.4.14.[illegible]) [illegible] bacterium [illegible]
C;Species: [illegible] bacterium meningosepticum
C;Date: 28-Oct-1996 #sequence_revision 1[illegible] Mar [illegible] #text_change [illegible]
C;Accession: S[illegible]
R;Kabashima, T.; Yoshida, T.; [illegible]
Arch. Biochem. Biophys. 320, 123-128, 1995
A;Title: Cloning, sequencing, and expression of the dipeptidyl-peptidase [illegible]
A;Reference number: S[illegible]; MUID:[illegible]; PMID:[illegible]
A;Accession: S[illegible]
A;Status: preliminary
A;Molecule type: DNA
A;Residues: [illegible] KAB[illegible]
A;Cross-references: EMBL:D42121; NID:[illegible]; PIDN:BAA0[illegible]; PID:[illegible]
C;Superfamily: dipeptidyl-peptidase IV
C;Keywords: dipeptidylpeptide hydrolase

  Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
  Best Local Similarity [illegible]; Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible] WSYYDN [illegible]

[illegible]

[illegible]
A;Molecule type: DNA
A;Residues: [illegible]
[illegible]
A;Experimental [illegible]
[illegible]
A;Molecule type: DNA
[illegible]
C;Genetics:
[illegible]
C;Superfamily: [illegible]
C;Keywords: dipeptidyl[illegible]

  Query Match [illegible]
  Best [illegible]
  Matches [illegible]

Qy [illegible]

Db [illegible]

[illegible]
CDH[illegible]
dipeptidyl-peptidase IV [illegible]
N;Alternate [illegible]
C;Species: [illegible]
C;Date: [illegible]
C;Accession: [illegible]
R;[illegible]
Biochem. [illegible]
A;Title: [illegible]
A;Reference [illegible]
A;Accession: [illegible]
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: [illegible]
R;[illegible]
J. [illegible]
A;Title: [illegible]
[illegible]
```

[illegible]

SUMMARIES

[illegible]

ALIGNMENTS

RESULT 1

[illegible]

RESULT 2

[illegible]

```
XX
PT   Novel polypeptides and polynucleotides useful for diagnosis,
PT   prevention, treating neural, immune system, [illegible] reproductive,
PT   pulmonary, cardiovascular, renal, proliferative disorders and [illegible]
PT   diseases
XX
PS   Claim 11; SEQ ID NO 1180; 1180pp; English.
XX
CC   The present invention relates to the isolation of novel human enzyme
CC   polypeptides, and the cDNA (AAS1079?-AAS1???) and genomic sequences
CC   encoding them. The enzyme polypeptides of the invention may comprise the
CC   functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC   isomerases or ligases. The sequences of the invention are useful in the
CC   diagnosis, treatment, prevention and/or [illegible]
CC   disorders including hyperproliferative disorders (e.g. cancer),
CC   immune deficiency disorders (e.g. AIDS) and immune disorders
CC   (e.g. arthritis), neurological disorder (e.g. Alzheimer's disease),
CC   metabolic disorders (e.g. phenylketonuria), inflammatory disorders
CC   (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis)
CC   [illegible] disorders (e.g. haemophilia), reproductive disorders
CC   (e.g. infertility) and infectious disorders (e.g. influenza). The
CC   polynucleotides of the invention can also be used in gene therapy.
CC   AAU2?? [illegible] AAU2??? represent the novel human enzyme polypeptides of the
CC   invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence [illegible] AA;

  Query Match             9?.2%;  Score ??;  DB 2?;  Length 1??;
  Best Local Similarity   88.?%;  Pred. No. [illegible];
  Matches    8;  Conservative    1;  Mismatches [illegible];  Indels [illegible]

Qy          HGWSY?YV? 9
              | |:|
Db      ?? HGWSY?IP? 5?


Search completed: February [illegible]
Job time : [illegible] secs
```

```
Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible]

Qy        [illegible] HWSYIIYL [illegible]

Db        [illegible] HWSYIIYL [illegible]


RESULT [illegible]
AAB24[illegible]
ID   AAB2[illegible] standard; peptide; [illegible] AA.
XX
AC   AAB2[illegible];
XX
DT   2[illegible] 200[illegible] (first entry)
XX
DE   Human dipeptidyl peptidase 9 (DPP9) peptide #1.
XX
KW   Human; dipeptidyl peptidase; DPP; neoplasia; type II diabetes; [illegible]
KW   autoimmunity; human immune deficiency virus; HIV infection; [illegible]
KW   graft rejection; antidiabetic; antiinflammatory; immunosuppressive;
KW   antiviral; enzyme.
XX
OS   Homo sapiens.
XX
PN   WO[illegible] A1.
XX
PD   [illegible] MAY [illegible]
XX
PF   [illegible] 200[illegible] 2001WO-AU0[illegible].
XX
PR   27[illegible] 2000AU-[illegible].
XX
PA   (UNSY ) UNIV SYDNEY.
XX
PI   Abbott CA, Gorrell MD;
XX
DR   WPI; 2[illegible] 454446/48.
XX
PT   New dipeptidyl peptidase (DPP) peptides, useful for screening
PT   inhibitors of DPP catalytic activity, which may be employed [illegible]
PT   [illegible] neoplasia, type II diabetes, cirrhosis, autoimmunity, graft
PT   rejection and HIV infection
XX
PS   Claim [illegible]; Page 28; [illegible] English.
XX
CC   The present invention relates to dipeptidyl peptidase (DPP) [illegible]
CC   polynucleotides encoding such proteins. The DPP peptides are useful [illegible]
CC   screening inhibitors of DPP catalytic activity. The inhibitors [illegible]
CC   for treating neoplasia, type II diabetes, cirrhosis, autoimmunity, graft
CC   rejection and HIV (human immune deficiency virus) infection. The present
CC   sequence is human DPP9 peptide.
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible] Score [illegible] DB [illegible] Length [illegible]
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible]

Qy        [illegible] HWSYIIYL [illegible]

Db        [illegible] HWSYIIYL [illegible]


RESULT 14
AAM[illegible]
ID   AAM[illegible] standard; Protein; [illegible] AA.
XX
AC   AAM[illegible];
XX
DT   [illegible] (first entry)
XX
DE   Human [illegible] peptide SEQ ID NO [illegible]
XX
```

```
KW   [illegible]
XX
OS   Homo sapiens.
XX
PN   [illegible] A1.
XX
PD   [illegible]
XX
PF   [illegible]
XX
PR   [illegible]
PR   [illegible] JAN [illegible]
PR   [illegible] MAY [illegible]
XX
PA   (HUMA-) [illegible]
XX
PI   [illegible]
PI   Wiseman [illegible]
XX
DR   WPI; [illegible]
DR   N-PSDB; AAS[illegible]
XX
PT   [illegible]
PT   [illegible]
XX
PS   [illegible]
XX
CC   [illegible]
XX
SQ   Sequence [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy        [illegible]

Db        [illegible]


RESULT [illegible]
AAG[illegible]
ID   AAG[illegible] standard; [illegible]
XX
AC   AAG[illegible];
XX
DT   [illegible]
XX
DE   New [illegible]
XX
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
XX
OS   Homo sapiens.
XX
PN   W[illegible]
XX
PD   [illegible]
```

```
PT   [illegible] infections, cancers, allergies, heart [illegible] disorders [illegible]
PT
XX
PS   [illegible] Disclosure; [illegible] English.
XX
CC   The present invention relates to the [illegible] novel human [illegible]
CC   proteases referred to as dipeptidyl peptidase IV (DPPIV) related
CC   proteins (DPRP). The dipeptidyl peptidase IV related proteins (DPRP)
CC   and nucleic acids encoding them are useful for treating infections
CC   such as fungal, bacterial, protozoan and viral infections, particularly
CC   infections caused by human immunodeficiency virus (HIV 1 or HIV 2),
CC   pain, diabetes, precocious puberty, infertility, obesity, anorexia,
CC   bulimia, Parkinson's disease, acute heart failure, hypotension,
CC   hypertension, urinary retention, osteoporosis, angina pectoris,
CC   stroke, ulcers, asthma, allergies, cancers, migraine, vomiting,
CC   psychotic and neurological disorders (e.g. anxiety, dementia, or
CC   schizophrenia), and dyskinesias. These may also be used in [illegible]
CC   therapeutic agents for the treatment of reproductive, inflammatory and
CC   metabolic disorders. [illegible] represent human DPRP proteins.
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible] Length [illegible]
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy        [illegible] WSY [illegible]

Db        [illegible] WSYPYL 744

RESULT [illegible]
AA [illegible]
ID   AA [illegible] standard; protein; [illegible] AA.
XX
AC   AA [illegible]
XX
DT   [illegible] (first entry)
XX
DE   Human protease PRISM protein sequence.
XX
KW   Human protease; PRISM; gastrointestinal; Crohn's disease; cancer;
KW   cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
KW   inflammatory disorder; acquired human deficiency syndrome; AIDS;
KW   cell proliferative disorder; developmental disorder; epilepsy;
KW   Duchenne muscular dystrophy; epithelial disorder; neurological [illegible]
KW   reproductive disorder; endometriosis.
XX
OS   Homo sapiens.
XX
PN   WO [illegible] A2.
XX
PD   [illegible]
XX
PF   [illegible] JUN [illegible] 2001 WO-US [illegible]
XX
PR   [illegible] JUN [illegible] 2000US [illegible]
PR   [illegible] JUN [illegible] 2000US [illegible]
PR   [illegible] JUN [illegible] 2000US [illegible]
PR   [illegible] JUL [illegible] 2000US [illegible]
PR   [illegible] JUL [illegible] 2000US [illegible]
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   [illegible] VS, [illegible] AP, [illegible]
PI   [illegible] AM, Baughn MR, [illegible] EA, [illegible] A, [illegible]
PI   [illegible] NK, Yao MG, [illegible] DAM, [illegible] Y, [illegible]
PI   [illegible] Y, [illegible] Y, [illegible] X, [illegible] Reddy R, [illegible]
PI   [illegible] EA;
XX
DR   WPI; [illegible]
DR   N-PSDB; AAK [illegible]
XX
```

```
PT   [illegible]
PT   [illegible]
PT   [illegible]
PT   [illegible]
XX
PS   [illegible]
XX
CC   The present [illegible] human [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   Crohn's disease [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
XX
SQ   Sequence [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy        [illegible] WSY [illegible]

Db        [illegible]

RESULT [illegible]
AA [illegible]
ID   AA [illegible]
XX
AC   AA [illegible]
XX
DT   [illegible]
XX
DE   [illegible]
XX
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
XX
OS   Homo sapiens.
XX
PN   WO [illegible]
XX
PD   [illegible]
XX
PF   [illegible]
XX
PR   [illegible]
XX
PA   [illegible]
XX
PI   [illegible]
XX
DR   WPI; [illegible]
DR   N-PSDB; [illegible]
XX
PT   [illegible]
PT   [illegible]
PT   [illegible]
XX
PS   [illegible]
XX
CC   [illegible]
```

XX
[illegible]
XX
CC This sequence represents human dipeptidyl aminopeptidase (DPP-, [illegible] has substrate specificity for H-Ala-Pro-pNA, [illegible] pNA and H-Arg-Pro-pNA. Therefore, it is a [illegible] peptidase and a dipeptidyl peptidase, because it is capable of hydrolysing the peptide bond [illegible] to Pro in each of these [illegible]. [illegible] is homologous with human DPPIV. [illegible] is useful for [illegible] substrates and for detecting an activated [illegible], which involves measuring the level of DPP8 gene expression [illegible]. The level of [illegible] expression is detected by detecting the amount of [illegible] RNA [illegible]. It is also useful for identifying a [illegible] capable of inhibiting the cleavage of the substrate by [illegible]. [illegible] identified as inhibiting [illegible] activity may be useful for treating [illegible], growth hormone deficiency, [illegible], non-insulin dependent diabetes [illegible] and [illegible] immunosuppressants.
XX
SQ Sequence [illegible] AA;

[illegible]

XX
[illegible]
XX
KW Human dipeptidyl peptidase [illegible] protein; [illegible] [illegible] HIV [illegible] [illegible] diabetes [illegible] [illegible]
XX
[illegible]
XX
[illegible]

XX
[illegible]
XX
[illegible]
XX
PI New dipeptidyl peptidase [illegible] type II diabetes [illegible] immunity [illegible]
XX
CC The present [illegible] relates to dipeptidyl peptidase [illegible] The [illegible] peptides [illegible] [illegible] type II diabetes [illegible] (human immune deficiency virus) [illegible] sequence [illegible] protein.
XX
SQ Sequence [illegible] AA;

Query Match [illegible] Score [illegible] Length [illegible]
Best Local Similarity [illegible] Pred. No. [illegible]
Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy [illegible]

Db [illegible]

RESULT [illegible]
[illegible]
ID [illegible] standard; protein; [illegible] AA.
XX
AC [illegible]
XX
DT [illegible]
XX
DE Human [illegible] protease [illegible]
XX
KW Human [illegible] dipeptidyl peptidase IV related protein; [illegible]
KW [illegible] HIV [illegible] HIV [illegible]
KW diabetes [illegible]
KW heart [illegible]
KW [illegible]
KW [illegible]
KW metastatic [illegible]
XX
OS Homo sapiens.
XX
PN W[illegible] A.
XX
[illegible]
XX
PF [illegible]
XX
PR [illegible]
XX
PA [illegible]
XX
PI [illegible]
XX
DR WPI; [illegible]
DR N-PSDB; AAK[illegible]
XX
PT New dipeptidyl peptidase IV related [illegible]
PT the protease [illegible] for treating [illegible] diabetes [illegible]

```
[illegible]
ID   [illegible] standard; protein; [illegible] AA.
XX
AC   [illegible]
XX
DT   [illegible] 2002  (first entry)
XX
DE   Human dipeptidyl peptidase [illegible].
XX
KW   Human dipeptidyl peptidase [illegible]; antiasthmatic; antiallergic;
KW   antiinflammatory.
XX
OS   Homo sapiens.
XX
PN   [illegible]
XX
PD   [illegible] 2002.
XX
PF   [illegible] 2001; 2001US-[illegible].
XX
PR   [illegible] 2001; 2001US-[illegible].
XX
PA   (BMFA-) BMFA CORP BV.
XX
PI   [illegible] E, Lacroix J, Monod M;
XX
DR   WPI; 2002-[illegible]/21.
XX
PT   Treating a patient for mucosal inflammation associated with rhinitis,
PT   sinusitis or both, by intranasally administering a peptidase that
PT   cleaves at Xaa-Pro sequences, to the patient
XX
PS   Disclosure; Column 23-24; [illegible]pp; English.
XX
CC   This invention relates to the treatment of a patient for mucosal
CC   inflammation associated with rhinitis or sinusitis, comprising
CC   intranasally administering a peptidase. The peptidase is [illegible]
CC   antiasthmatic, antiallergic and antiinflammatory in its action.
CC   The peptidase cleaves at Xaa-Pro sequences and is useful for treating
CC   a patient for mucosal inflammation associated with rhinitis or
CC   sinusitis, which is the result of allergies or asthma. This
CC   sequence represents human dipeptidyl peptidase [illegible].
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
  Best Local Similarity [illegible]; Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy   [illegible] WSYIYL [illegible]

Db   [illegible] WSYIYL 172

RESULT [illegible]
AAB[illegible]
ID   AAB[illegible] standard; Protein; [illegible] AA.
XX
AC   AAB[illegible]
XX
DT   [illegible] 2001  (first entry)
XX
DE   Human [illegible] Val-[illegible].
XX
KW   Human; dipeptidyl aminopeptidase; [illegible] peptidase;
KW   dipeptidyl peptidase; DPP IV; [illegible] diarrhoea;
KW   growth hormone deficiency; glucose [illegible]
KW   non-insulin dependent diabetes mellitus; [illegible]
KW   immune suppression.
XX
OS   Homo sapiens.
XX
PN   W[illegible]-A1.
```

```
XX
     [illegible]
XX
PD   [illegible]
XX
PF   [illegible]
PR   [illegible]
XX
PA   [illegible]
XX
PI   [illegible]
XX
DR   WPI; [illegible]
DR   N-PSDB; [illegible]
XX
PT   New human [illegible]
PT   [illegible]
PT   [illegible]
XX
PS   [illegible]
XX
CC   [illegible]
CC   dipeptidyl [illegible]
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy   [illegible] WSY[illegible]

Db   [illegible] WSY[illegible]

RESULT [illegible]
AAB[illegible]
ID   AAB[illegible]
XX
AC   AAB[illegible]
XX
DT   [illegible]
XX
DE   Human [illegible]
XX
KW   Human [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
XX
OS   [illegible]
XX
PN   [illegible]
PN   [illegible]
```

[illegible]

[illegible]

[illegible]